

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number  
**WO 00/73478 A2**

- (51) International Patent Classification<sup>1</sup>: C12N 15/861, 15/864, 15/10, A61K 48/00 (74) Agent: ADRIANO, Sarah, B.; Mandel & Adriano, Suite 60, 35 No. Arroyo Parkway, Pasadena, CA 91103 (US).
- (21) International Application Number: PCT/US00/15442 (81) Designated States (*national*): AE, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (22) International Filing Date: 1 June 2000 (01.06.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/137,213 1 June 1999 (01.06.1999) US  
60/161,097 22 October 1999 (22.10.1999) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): UNIVERSITY OF WASHINGTON [US/US]; Suite 200, 1107 N.E. 45th Street, Seattle, WA 98105-4631 (US).
- (72) Inventors; and  
(75) Inventors/Applicants (*for US only*): LIEBER, André [DE/US]; 7251 29th Avenue, Seattle, WA 98115 (US). SHAYAKHMETOV, Dmitry [RU/US]; 6500 25th Avenue, Seattle, WA 98115 (US). FARRAR, Denise [US/US]; 1324 W. Emerson Street, Seattle, WA 98119 (US). PAPAYANNOPOULOU, Thalia [US/US]; 702 35th Avenue, Seattle, WA 98122 (US).

**Published:**

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/73478 A2

(54) Title: RECOMBINANT ADENOVIRAL VECTORS FOR CELL SPECIFIC INFECTION AND GENOME INTEGRATION AND EXPRESSING CHIMERIC FIBER PROTEINS

(57) Abstract: The present invention provides for novel chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_{3/5}$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

5      **RECOMBINANT ADENOVIRAL VECTORS FOR CELL SPECIFIC  
INFECTION AND GENOME INTEGRATION AND EXPRESSING CHIMERIC  
FIBER PROTEINS**

This invention was made, at least in part, with funding from the National Institutes of Health (Grant Nos. R01 CA 80192-01 and R21 DK 55590-01). Accordingly, the United  
10      States Government has certain rights in this invention.

**FIELD OF THE INVENTION**

This invention relates to the field of gene therapy, and in particular, to novel adenovirus  
15      (Ad) vectors that selectively infect cells for gene therapy, and to Ad vectors containing modifications of the fiber protein to allow retargeting of any adenovirus serotype.

**BACKGROUND OF THE INVENTION**

20      Gene transfer vectors require the efficient transduction of target cells, stable association with the host genome, and adequate transgene expression in the appropriate target cell, without associated toxic or immunological side effects,. Currently available viral vector systems, including recombinant retroviruses, adenoviruses and adeno-associated viruses, are not suitable for efficient gene transfer into many cell types. Retroviral vectors require  
25      cell division for stable integration. Recombinant adenoviruses are not able to infect many cell types important for gene therapy, including hematopoietic stem cells, monocytes, T- and B-lymphocytes. Moreover, recombinant adeno-associated vectors (AAV) integrate with a low frequency.

30      First generation adenoviruses have a number of properties that make them an attractive vehicle for gene transfer (Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205). These include the ability to produce purified virus at high titers in concert with highly efficient gene transfer of up to 8 kb long expression cassettes into a large variety of cell



types *in vivo*, including non-dividing cells. Limitations of first generation adenoviruses include the development of immune responses to expressed viral proteins resulting in toxicity and virus clearance. The episomal status of adenoviral DNA within transduced cells is another limitation of first generation Ad vectors. Stable integration of adenovirus DNA into the host genome is reported only for wild-type forms of specific subtypes and appears not to occur in a detectable manner with E1/E3-deleted Ad 5 (adenovirus serotype 5) vectors widely used for gene transfer *in vitro* and *in vivo* [Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205].

10 Recombinant AAV vectors (rAAV) integrate with a low frequency (about 1 out of 20,000 genomes) randomly as cocatemers into the host genome (Rutledge, E.A.; Russel, D.W. 1997 *J. Virology*, 71, 8429-8436). The presence of two AAV inverted terminal repeats (ITRs) and as yet unknown host cellular factors seem to be the only requirement for vector integration (Xiao, X., et al, 1997, *J. Virology*, 71, 941-948; Balague, C., et al. 15 1997, *J. Virology*, 71, 3299-3306; Yang, C.C. 1997, *J. Virology*, 71, 9231-9247). In the presence of the large AAV Rep proteins, AAV integrates preferentially into a specific site at human chromosome 19, called AAVS1 (Berns, K.I., 1996. *Fields Virology*, Fields, B.N. et al. (ed) Vol. 2, Lippincott-Raven, Philadelphia, PA, 2173-2220). The AAV capsid is formed by three coat proteins (VP1-3), which interact with specific heparin 20 sulfates on the cell surface and probably with specific receptor(s). However, many cell types, including hematopoietic stem cells, lack these structures so that rAAV vectors based on AAV2 cannot infect or transduce these cells (Malik P. et al., 1997, *J. Virology*, 71, 1776-1783; Quing, K.Y., et al. 1998, *J. Virology*, 72, 1593-1599). Other disadvantages of rAAV vectors include the limited insert size (4.5-5kb) that can be 25 accommodated in rAAV vectors lacking all viral genes and low transducing titers of rAAV preparations.

Adenovirus infection is initiated by attaching to the cell surface of Ad 5 via its fiber protein (for a review, see Shenk, T. 1996 *Fields Virology*, Vol. 2, Fields, B.N. et al. (ed) 30 Vol. 2, Lippincott-Raven, Philadelphia, PA, 2111-2148). The distal, C-terminal domain of the trimeric fiber molecule terminates in a knob, which binds to a specific cellular

receptor identified recently as the coxackie-adenovirus receptor (CAR) (Bergelson, J.M. et al. *Science*, 275, 1320-1323). After binding, in an event independent of virus attachment, Arg-Gly-Asp (RGD) motifs in the penton base interact with cellular integrins of the  $\alpha 3$  and  $\beta 5$  types. This interaction triggers cellular internalization whereby the virion achieves localization within the endosome. The endosomal membrane is lysed in a process mediated by the penton base, releasing the contents of the endosome to the cytoplasm. During these processes, the virion is gradually uncoated and the adenoviral DNA is transported to the nucleus where replication takes place. The terminal protein, which is covalently attached to the viral genome and the core protein V that is localized on the surface of the cores have nuclear localization signals (NLSs) (van der Vliet, B. 1995, *The Molecular Repertoire of Adenoviruses*, Vol. 2, Doerfler, W. and Boehm, P.(ed.), Springer Verlag, Berlin, 1-31). These NLSs play a crucial role in directing the adenoviral genome to the nucleus and probably represent the structural elements which allow adenovirus to transduce non-dividing cells. When the double-stranded, linear DNA reaches the nucleus, it binds to the nuclear matrix through its terminal protein.

Since the cell types that can be infected with Ad5 or Ad2 vectors are restricted by the presence of CAR and specific integrins, attempts have been made to widen the tropism of Ad vectors. Genetic modification of adenovirus coat proteins to target novel cell surface receptors have been reported for the fiber (Krasnykh, V. et al. 1998 *J. Virology*, 72, 1844-1852, Krasnykh, V. et al. 1996 *J. Virology*, 70, 6839-6846, Stevenson, S.D., et al. 1997, *J. Virology*, 71, 4782-4790), penton base (Wickham, T.J., et al. 1996, *J. Virology*, 70, 6831-6838; Wickham, T.J., et al. 1995, *Gene Therapy*, 69, 750-756), and hexon proteins (Crompton, J., et al. 1994, *J. Gen. Virol.* 75, 133-139). The most promising modification seems to be the functional modification of the fiber protein or more specifically of the fiber knob as the moiety, which mediates the primary attachment. Two groups have reported the generation of fibers consisting of the Ad5 tail/shaft and the knob domain of Ad3 (Krasnykh, V. et al. 1996 *supra*, Stevenson, S.D., et al. 1997, *supra*). Recently, recombinant adenoviruses with fibers containing C-terminal poly-lysine, gastrin-releasing peptide, somatostatin, E-selectin-binding peptide, or oligo-Histidines were produced in order to change the native tropism of Ad5. Krasnikh et al. found (Krasnykh, V. et al.

1998 *supra*) that heterologous peptide ligands could be inserted into the H1 loop of the fiber knob domain without affecting the biological function of the fiber. Based on studies with other Ad serotypes, it appears that the length of the fiber shaft is a critical element, determining the efficiency of interaction with cell surface integrins and the internalization process. Thus far, there is no reported data demonstrating successful retargeting of Ad5 vectors for a specific cell type.

Therefore, there is a present need for an improved adenovirus vector which can be targeted efficiently to a variety of cell types and tissues and remain stably integrated in the host genome with minimal antigenicity to the host. The present invention discloses novel chimeric adenoviral (Ad) Ad-AAV vectors, which express a modified fiber protein on their capsid, for specifically targeting the vector. Methods of making, uses and advantages of these vectors are described. In addition, the alteration described for the knob and shaft domains of the fiber protein provide a novel approach to retarget any adenovirus serotype for cell specific infection.

#### **SUMMARY OF THE INVENTION**

The present invention provides for novel chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_{3/5}$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

The recombinant adenovirus vectors of the invention (Example I) provide a novel design that allows for the easy production and delivery of a "gutless" adenoviral vector with the added advantage of stable integration of the transgene into the host genome of different cell type. The adenoviral vector described is devoid of all adenoviral sequences except for the 5' and 3' cis elements necessary for replication and virion encapsidation. The adenovirus-associated virus sequences of the invention comprising the 5' (right) and 3' (left) inverted terminal repeats (ITRs) flank the transgene gene cassette such that they

direct homologous recombination during viral replication and viral integration into the host genome. In one embodiment AAV-ITR flanking sequences are used. The vector also contains a selected transgene(s) operably linked to a selected regulatory element and a polydenylation stop signal, which is in turn flanked by the flanking sequences described  
5 above. The selected transgene(s) can be linked under the same regulatory elements or under separate regulatory elements in the same orientation or in opposite orientations with respect to each other. The selected transgene(s) are any gene or genes which are expressed in a host cell or tissue for therapeutic, reporter or selection purposes. This vector is characterized by high titer transgene delivery to a host cell and the ability to  
10 stably integrate the transgene into the host genome. Also provided is a method to improve the integration frequency and site specific integration by incorporating an AAV rep protein into the recombinant hybrid vector.

The invention also provides chimeric fiber proteins (Example II), which includes  
15 naturally occurring fiber proteins in which a portion or portions of the sequence are modified to alter cell or tissue specificity of infection. Altered fiber protein sequences can include fiber protein domains (the knob domain, the shaft domain, and the tail domain) from other or the same adenovirus serotypes or from randomly selected peptides. A chimeric fiber protein can be entirely composed of non-naturally occurring  
20 sequences. The invention further relates to nucleic acid sequences encoding the chimeric fiber proteins. These nucleic acid sequences can be naturally occurring, a mixture of naturally occurring and non-naturally occurring sequences, or entirely non-naturally occurring sequences.

25 The heterologous fiber protein sequences described herein can be inserted into any adenovirus based vector which contains a capsid, rendering the virus capable of specifically infecting a given cell or tissue. Adenoviral vectors having such a heterologous fiber sequence can be used to direct gene transfer into desired cells. For stable integration of the transgene cassette into the host genome, the chimeric Ad-AAV  
30 vector described in the invention is the preferred vector of use.

The invention also includes a library of adenoviruses displaying random peptides in their fiber knobs can be used as ligands to screen for an adenovirus variant with tropism to a particular cell type *in vitro* and *in vivo*.

- 5 The chimeric Ad- vectors described herein include the Ad.AAV genome with a modified fiber protein expressed on its capsid. These chimeric vectors are designed to infect a wide variety of cells, in particular, the cells which can only be poorly transduced by the commonly used retroviral, AAV and adenoviral vectors. These cells include, but not limited to, hematopoietic stem cells, lung epithelial cells, dendritic cells, lymphoblastoid  
10 cells, and endothelial cells. Hematopoietic stem cells such as CD34+ cells can be targeted for gene therapy of sickle cell anemia and thalasemia using the vector described herein. The chimeric Ad-AAV vector capable of transducing genes into endothelial cells can be used in gene therapy for vascular diseases such as atherosclerosis or restinosis after coronary artery surgery.

15

#### **BRIEF DESCRIPTION OF THE FIGURES**

Figs. 1A-1C display a proposed mechanism for forming of  $\Delta$ Ad.AAV1 genome.

- 20 Figs. 2A and 2B show electron photomicrographs of hybrid virus particles: Fig. 2A shows Ad.AAV1 and Fig. 2B shows  $\Delta$ Ad.AAV1.

- Fig. 3 illustrates analysis of  $\Delta$ Ad.AAV1 genomes after transduction of SKHep1 cells. Pulse field gel eletrophoresis (PFGE).  $1 \times 10^6$  control Sk Hep1 cells (SKHep1) (lanes 1-3, 5, 9). SKHep1 cells from G418 resistant pools ( $\Delta$ Ad.AAV1) (infected with  $\Delta$ Ad.AAV1 and selected for 4 weeks) (lanes 6-8, 10-12), or SKHep1 cells collected at 3 days after  
25 infection with 2000 genomes Ad.AAV1 (Ad) lanes 4, 13) are sealed in agarose plaques, lysed *in situ* and subjected to PFGE with or without prior digestion with restriction endonucleases. Southern Blot is performed with a SEAP specific probe. U = undigested,  
30 P = digested with PI- Scl, I = I-CeuI, E = EcoREI.

Figs. 4A and 4B show response of K562 and CD34+ cells respectively after infection with  $\Delta$ Ad.AAVBG. Cells are incubated for 6 hours with virus under agitation. At day 3 after infection, transduction frequency is calculated based on the number of X-Gal positive cells. Viability is tested by trypan blue exclusion. N = 3, SEM < 10%.

5

Fig. 5 shows Rep expression in SKHep1 and 293 cells after plasmid transfection.  $5 \times 10^5$  cells are transfected with pAAV/Ad, pRSVrep, or pPGKrep by Ca-phosphate co-precipitation. Three days after transfection, cells are harvested. Lysates are separated on a 10% PA gel, followed by Western Blot with Rep specific antibodies (03-65169),  
10 American Research Products), and developed with ECL (Amersham).

Fig. 6 shows detection of vector integration into AAVS1 by PFGE.

Fig. 7 shows strategy for creating an  $\Delta$ Ad.AAV hybrid vector capable of site-specific  
15 integration. Arrows indicate promoters, (PA) = polyadenylation signal.  $\Psi$  = adenoviral packaging signal.

Fig. 8A-8B shows vectors for transduction studies with SNori as expression unit and analysis of vector integration on genomic DNA from a small cell number. Analogous  
20 vector sets can be generated with  $\beta$ -galactosidase (BG) or green fluorescence protein (GFP) as reporter genes.

Fig. 9 shows strategy for substituting the Ad5 fiber sequence by the heterologous fiber X genes using recombination in *E. coli*.

25

Fig. 10 shows the expression of CAR and  $\alpha_v$ -integrins on test cells. For flow cytometry analysis, HeLa, CHO, K562, and CD34+ cells were incubated with monoclonal anti-CAR (RmcB, 1:400 dilution) or anti- $\alpha_v$ -integrin antibodies (L230, 1:30 dilution). As a negative control, cells were incubated with an irrelevant mouse monoclonal antibody (anti-BrdU,  
30 1:100 dilution). The binding of primary antibodies was developed with anti-mouse IgG-

FITC labeled conjugates (1:100 dilution). Data shown represent the average results of quadruplicate analyses performed on  $10^4$  cells.

Fig. 11 shows the electron microscopy of adenovirus particles. Purified particles from Ad5, 9, and 35 were negative contrast stained and analyzed at a magnification of 85,000x. Defective particles are highlighted by arrows.

Fig. 12 shows the analysis of attachment and internalization of different serotypes to CHO, HeLa, K562, and CD34+ cells. Equal amounts of [ $^3$ H]-thymidine-labeled virions of Ads 3, 4, 5, 9, 35, and 41 (measured by OD<sub>260</sub>, and equivalent to an MOI of 400 pfu per cell for Ad5) were incubated for one hour on ice as described in Materials and Methods. Cells were then washed, and the number of labeled virions bound per cell was determined. For internalization studies, viruses were first allowed to attach to cells for 1 h on ice. Then, unbound viral particles were washed out. Cells were then incubated at 37°C for 30 min followed by treatment with trypsin-EDTA and washing to remove uninternalized viral particles. The data were obtained from two to four independent experiments performed in triplicate. Note the different scale on the Y-axes for CD34+ cells.

Figs. 13A-13C show attachment and internalization of different adenovirus serotypes to HeLa, CHO and 293 cells respectively.

Figs. 14A and 14B show attachment and internalization of different adenovirus serotypes to CD34+ and K-562 cells respectively.

Fig. 15A-15C shows the analysis of viral replication in K562 and CD34+ cells by Southern blot analysis of methylated viral DNA. Replication studies were performed with  $1 \times 10^5$  K562 cells (A) or CD34+ cells (B), infected with methylated Ad5, Ad9 or Ad35. The lane labeled as "load" represents DNA that was extracted from the media/cell mixture immediately after adding the indicated viral dose to cells. The intensities of bands corresponding to methylated and un-methylated viral DNA indicate that ~85% of

the input virus was methylated. To quantify adsorption and internalization, DNA analysis was performed after prior incubation of virus with cells at 0°C (adsorption) or 37°C (internalization). For dose dependent replication studies, the indicated viral dose (expressed as the number of genomes) was added to the cells, and cellular genomic DNA together with viral DNA was extracted 16 hours or 36 hours post-infection for K562 and CD34+cells, respectively. Identical amounts of sample DNA were analyzed by Southern blot. For quantification purposes, Ad9 replication was analyzed together with Ad5 using an Ad5/9 chimeric probe that hybridizes with DNA of both serotypes (C). The analysis of Ad5 versus Ad35 replication was performed with the corresponding Ad5/35 chimeric probe. Since separate hybridizations with both Ad5/35 and Ad5/9 probes gave identical signal intensities for Ad5 DNA only one panel is shown for Ad5 replication in test cells. To produce distinguishable fragments specific for the methylated or non-methylated status of viral genomes, Ad5 DNA was digested with *Xho* I, while Ad9 and Ad35 DNA was digested with *Xho* I and *Hind* III. The bands specific for methylated (not-replicated) viral DNA were ~12kb for Ad9, 35kb for Ad5, and ~12kb for Ad35. The fragments specific for non-methylated DNA were 5.8kb for Ad9, 6.1kb for Ad5, and 9.5kb for Ad35. Chimeric Ad5/9 and Ad5/35 DNA fragments (1.8kb) were used as quantification standards and applied onto gel together with digested viral/cellular DNA (shown on the left part of the figures).

20

Fig. 16A-16B shows the structure of Ad5GFP and chimeric Ad5GFP/F35 vectors. A) Schematic diagram of the original E1/E3 deleted Ad5-based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes. A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The

30



Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal. The region of Ad5GFP/F35 encoding for chimeric fiber was completely sequenced with Ad5 specific primers (see Material and Methods). B) Restriction analysis of viral genomes. Viral DNA was isolated from purified Ad5GFP and Ad5GFP/F35 particles as described elsewhere. One microgram of DNA was digested with *Hind* III or *Kpn* I and separated in ethidium bromide stained agarose gels (left panel) which were subsequently blotted and analyzed by Southern blot with an Ad5 E4 specific probe (nt 32,7775-33,651) (right panel). Specific patterns, designating the correct structure for both viral vectors were detected. The *Hind* III fragments specific for Ad5GFP and Ad5GFP/F35 were 2.9kb and 4.9kb, respectively. The *Kpn* I fragment that confirmed the correct Ad5GFP/F35 structure was 1.6kb compared to a 7.6kb Ad5GFP fragment. M - 1kb ladder (Gibco-BRL, Grand Island, NY).

Fig. 17 shows the generation of  $\Delta$ Ad.AAV genomes by recombination between inverted homology regions. Recombination between two inverted repeats (IR) present in one Ad.AAV vector. The first-generation Ad.AAV vector (~34kb) contains two 1.2kb inverted homology elements flanking the transgene cassette. One AAV-ITR is inserted between the Ad packaging signal ( $\psi$ ) and the left IR. During Ad replication, recombination between the Irs mediates the formation of the  $\Delta$ Ad.AAV genomes with the transgene flanked by Irs, AAVITRs, Ad packaging signals, and Ad ITRs. These genomes are efficiently packaged into Ad capsids.

Fig. 18 shows the structure of Ad5/11, Ad5/35. Schematic diagram of the original E1/E3 deleted Ad50based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes.

A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal.

5 Fig. 19 shows the cross-competition for attachment and internalization of labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled viruses, and with anti-CAR or anti- $\alpha_v$ -integrins Mab. (A) For attachment studies,  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled competitor virus at 4°C for 1 h. Then, equal amounts of [ $^3$ H]-labeled viruses, at a dose equivalent to an MOI of 100 pfu per cell  
10 determined for Ad5GFP, were added to cells followed by incubation at 4°C for 1 h. Cells were then washed with ice-cold PBS, pelleted and the percentage of attached virus (cell-associated counts per minute) was determined. For analysis of cross-competition for internalization, cells were pre-incubated with a 100-fold excess of competitor virus at 37°C for 30 min before labeled virus was added. After an additional incubation at 37°C  
15 for 30 min, cells were treated with trypsin-EDTA for 5 min at 37°C, washed with ice-cold PBS, pelleted, and the percentage of internalized virus was determined. For controls, cells were incubated with labeled viruses without any competitors. Preliminary experiments had shown that the conditions chosen for competition studies allowed for saturation in attachment/internalization on K562 cells for all unlabeled competitors. (B)  $10^5$  K562 cells  
20 were pre-incubated for 1 hour at 4°C with anti-CAR MAb (RmcB, diluted 1:100) or with anti- $\alpha_v$ -integrin MAb (L230, diluted 1:30), followed by incubation with labeled viruses according to the protocols for attachment or for internalization as described above. For each particular serotype, the percentage of attached/internalized virus was compared to the control settings, where cells were preincubated under the same conditions with a  
25 1:100 dilution of an irrelevant antibody (anti-BrdU Mab) before addition of the labeled virus. Note that the specific competitors but not the corresponding controls significantly inhibited Ad5 internalization to a degree that is in agreement with published data (59).  
N>/=4

30 Fig. 20. Cross-competition for attachment and internalization of [ $^3$ H]-labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled Ad3 virus (A), and of [ $^3$ H]-

labeled Ad3 virions with unlabeled viruses (B).  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled viral particles according to attachment or internalization protocols described for Fig.6. Equal amounts of [ $^3$ H]-labeled Ad5GFP, Ad5GFP/F35, or Ad35 (A) or [ $^3$ H]-labeled Ad3 (B) were added to cells at a dose equivalent to an MOI of 100 pfu per cell for Ad5GFP. In control settings, cells were incubated with labeled viruses without any competitors. N=4.

Fig. 21 shows the transduction of CD34+, K562, and HeLa cells with Ad5GFP and chimeric Ad5GFP/F35 vectors.  $1 \times 10^5$  cells were infected with different MOIs (pfu/cell) of viruses in 100  $\mu$ l of media for 6 hours at 37°C. Virus containing media was then removed, and the cells were resuspended in fresh media followed by incubation for 18 h at 37°C. The percentage of GFP expressing cells was determined by flow cytometry. N=3

Fig. 22 shows the distribution of GFP-positive cells in subpopulations of human CD34+ cells expressing CAR or  $\alpha_v$ -integrins.  $1 \times 10^5$  CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu/cell as described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CAR (1:100 final dilution) or anti- $\alpha_v$ -integrin (1:30 final dilution) primary MABs for 1 h at 37°C. Binding of primary antibodies was developed with anti-mouse IgG-PE labeled secondary MABs (1:100 final dilution) at 4°C for 30 min. For each variant,  $10^4$  cells were analyzed by flow-cytometry. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The data shown were representative for three independent experiments.

Fig. 23A-23B shows the distribution of GFP-positive cells in a subpopulation of human CD34+ cells, expressing CD34 and CD117 (c-kit). (A) Co-localization of GFP expression with CD34 or CD117: CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell under the conditions described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CD34 PE-conjugated MABs (final dilution 1:2) or with anti-CD117 PE-conjugated MABs (final dilution 1:5)

for 30 min on ice, and  $10^4$  cells per variant were subjected to two-color flow cytometry analysis. For negative control staining, no antibodies were added to the cells before analysis. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The experiment was performed two times in triplicates, and typically obtained results are shown. The SEM was less than 10% of the statistical average. (B) Transduction of CD34+/CD117+ cells with Ad5GFP and chimeric Ad5GFP/F35 virus vectors: CD34+ cells, cultured overnight before staining in media without SCF, were incubated with PE-labeled anti-CD117 MAb for 30 min on ice. The fraction of CD117-positive cells was sorted by FACS. More than 97% of sorted cells were positive for CD117.  $1 \times 10^5$  CD117+/CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell, as for Fig.8. Twenty-four hours post infection, the percentage of GFP positive was determined by flow cytometry. For mock infection, CD117+/CD34+ cells were incubated with virus dilution buffer only. The infections were done in triplicates, and the average percentage of GFP-expressing cells is indicated on the corresponding histogram. The SEM was less than 10% of the statistical average.

Fig. 24 shows the southern analysis of viral genomes in GFP-positive and GFP-negative fractions of CD34+ cells infected with the Ad5GFP and chimeric Ad5GFP/F35 vectors. CD34+ cells were infected with viruses at an MOI of 100 as described for Fig.21. Twenty four-hours post infection, cells were sorted by FACS for GFP positive and GFP negative fractions.  $10^5$  cells from each fraction were used to isolate genomic DNA together with viral DNA. Before cell lysis, a rigorous treatment with trypsin and DNase followed by washing was performed to exclude that genomic DNA samples were contaminated by extracellular viral DNA. A) The upper panel shows the ethidium bromide stained 1% agarose gel before blotting demonstrating that similar amounts of genomic DNA were loaded. This amount corresponded to DNA isolated from ~25,000 GFP+ or GFP- cells. The lane labeled Aload@ represents viral DNA purified from Ad5GFP or Ad5GFP/F35 virions mixed with pBluescript plasmid DNA (Stratagene) as a carrier and applied on a

gel at the amount that was actually used to infect 25,000 cells. As a concentration standard, a serial dilution of Ad5GFP genomes was loaded on the gel (left side). For Southern analysis (lower panel), an 8kb-long HindIII fragment corresponding to the E2 region of Ad5 was used as a labeled probe. Hybridized filters were subjected to  
5 PhosphorImager analysis and then exposed to Kodak-X-OMAT film for 48 h at B70°C. The cellular/viral genomic DNA is indicated by an arrow. (B) To detect Ad5GFP genomes in transduced cells, PCR amplification followed by Southern blot hybridization was performed on the same samples that were used for quantitative Southern blot hybridization in (A). DNA purified from ~2,500 cells was subjected to PCR (95°C B1min,  
10 53°C-1min, 72°C B1min, 20 cycles with primers Ad5-F1 and Ad5-R1). One fifth of the PCR reaction was subjected to agarose gel electrophoresis (upper panel). A 0.9 kb-long DNA fragment, specific to the E4 region of Ad5 was detected for transduced Ad5GFP/F35 genomes. DNA then was blotted onto Nybond-N+ membrane and Southern blot hybridization (lower panel) with an Ad5 E4 specific DNA probe was performed. In  
15 addition to the 0.9kb DNA fragment, the PCR primers generated a smaller 0.5 kb-long fragment that also hybridized with with the E4 region probe.

Fig. 25 shows the role of fiber shaft length in Ad infection strategies. CAR binding (Ad5 and Ad9) variants and Ad35, which interacts with a non-CAR receptor were analyzed on  
20 CAR expressing cells (293, Y79) and K562 cells which do not express significant CAR amounts. All vectors contain a GFP expression cassette packaged into an Ad5 capsid with modified fibers.

Fig. 26 shows the tertiary structure of Ad5 knob: localization of CAR binding sites, H-I  
25 loop and G-H loop.

Fig. 27 shows the substitution of the G-H loop with heterologous peptides.

Fig. 28 shows the attachment and internalization of metabolically labeled serotypes with  
30 human cell lines.

Figs. 29A-29D shows the generation of Rep78 expressing Ad vectors by recombination between two vectors. (A) The same strategy outlined in Fig. 15 was employed for vectors with rep 78 as a transgene. The Ad5'rep vector also contained the ApoEhAAT promoter shielded by an HS-4 insulator. The region of homology between the two fragments of the rep78 gene was 658nt in length. The Rep78 ORF was deleted for the p5 promoter. The internal Rep 40/52 start codon (at position 993) was mutated to abolish production of the small Rep proteins. Furthermore, the splice site at nt 1905 was deleted eliminating production of Rep68. The individual expression of Rep 78 was demonstrated. (B) Formation of  $\Delta$ Ad.rep78 genomes. The expected 5.8kb  $\Delta$ Adrep78 genome was only observed upon coinfection of both Ad5'rep and Ad3'rep into 293 cells as demonstrated by Southern. (C) Southern blot analysis for rescue of the recombinant AAV genome from plasmid DNA by Rep78 expressed from pCMVrep78 and  $\Delta$ Ad.rep78. The expected rescue product is 3.8kb (R-plasmid). (D) Southern blot analysis for rescue of the recombinant AAV genome from Ad.AAV viral vector genomes.

15

### **DETAILED DESCRIPTION OF THE INVENTION**

All scientific and technical terms used in this application have meanings commonly used in the art unless otherwise specified. As used in in the present invention, the following words or phrases have the meanings specified.

The term vector includes, but is not limited to, plasmids, cosmids, phagemids, and artificial chromosomes. The vector sequence may be designated as the viral "base vector" sequence. The base vector sequence is dependent upon the particular type of virus and serotype that the base vector sequence was derived from. The base vector sequence may be linked to non-vector or transgene sequences (e.g., heterologous sequences).

The transgene sequences may include sequences that confer compatibility with prokaryote or eukaryote host cells, where compatibility relates to vector replication

within a host cell. Accordingly, the transgene sequence may be a replicon sequence that directs replication of the vector within the host cell, resulting in an autonomously replicating vector. Alternatively, the transgene sequence may permit vector replication that is dependent upon the host cell's replication machinery.

5

The base vector sequences may be operatively linked to a transgene sequence that encodes a gene product, such as a polypeptide, rRNA, or tRNA. For example, the transgene may encode a polypeptide such as a viral capsid protein, or a viral fiber protein. The transgene may be derived from the same or different serotype as the base vector sequence.

10

Another example of a transgene includes a reporter gene that encodes a gene product that can be used as a selectable marker, such as drug resistance or a colorimetric marker. The reporter gene may encode a gene product which can be readily detected by, for example, a visual microscopic, immunochemical, or enzymatic assay. The preferred reporter gene encodes a gene product that can be detected by a non-destructive method that does not destroy the cell that expresses the reporter gene.

15

A therapeutic gene is another example of a transgene. A therapeutic gene encodes a gene product (e.g., polypeptide or RNA) which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or the tissue or the organ or the organism containing the host cell. The therapeutic benefit may result from modifying a function of a gene in the host genome or from the additional function provided by the therapeutic protein, polypeptide or RNA.

20

The base vector sequence may be linked to a transgene sequence that is a regulatory element, such as a promoter, enhancer, transcription termination signals, polyadenylation sequences. The regulatory element may direct expression of the transgene sequence that encodes a gene product by direct transcription or translation.

25

The regulatory element may regulate the amount or timing of expression of the transgene

30

sequence. The regulatory element may direct expression of the transgene in certain host cells or tissues (e.g., host-specific or tissue-specific expression).

5 The base vector sequence may be linked to a transgene sequence that permits the vector, to integrate into another nucleotide sequence. The integration sequence may direct integration of the whole vector or portions of the vector. The integration sequence may or may not be related to the base vector sequence. For example the integration and base vector sequences may be from the same or different viral serotype. The integration sequence may be inverted repeat sequences (ITRs) from adenovirus (Ad), adenovirus-  
10 associated virus (AAV), or HIV.

The base vector sequence may be linked to a transgene sequence that directs homologous recombination of the vector into the genome of a host cell. Such transgene sequences may or may not be from the same viral serotype as the base vector sequence.

15

The vector may be used to transport the heterologous sequence into a host cell or into a host cell's genome.

The vector may comprise multiple endonuclease restriction sites that enable convenient  
20 insertion of exogenous DNA sequences.

The term "hybrid vector" as used in the invention refers to a vector which comprises a nucleic acid sequence combined from two different viruses (e.g. Adenovirus and AAV).

25 "Chimeric vector" refers to a vector which contains nucleic acid sequences that are unnatural to the base vector (i.e. sequences not occurring naturally or sequences not in their natural background including heterologous sequences). A chimeric vector as used in the invention may also be a hybrid vector. An example of a chimeric vector is Ad.AAV expressing a modified fiber protein on its capsid.

30



The term "transduction" or "infection" refers to a method of introducing viral DNA within a virus particle into a host cell. The viral DNA herein is in the form of recombinant virus, which is generated by linking a segment of DNA of interest into the viral genome in such a way that the gene can be expressed as a functional protein.

5

The term "transfection" refers to a method of introducing a DNA fragment into a host cell.

The term "heterologous" as used herein means that a nucleic acid or peptide sequence is placed in a context that is not endogenous to the base adenovirus vector or to a transduced cell. For example, a peptide sequence can be transferred from a protein to another protein, the resulting protein is referred to herein as heterologous protein. A chimeric fiber protein, (e.g., a serotype 5 tail domain and a serotype 35 shaft and knob domain) is considered a "heterologous" to the Ad 5 vector. The term also includes nucleic acids (e.g. coding sequences) from one strain or serotype of adenovirus introduced into a different strain or serotype of adenovirus.

The term "regulatory elements" is intended to include promoters, enhancers, transcription termination signals, polyadenylation sequences, and other expression control sequences. Regulatory elements referred to in the invention include but are not limited to, those which direct expression of nucleic acid sequence only in certain host cells (e.g. tissue specific regulatory sequences).

The term "operably linked" indicates that a polynucleotide sequence (e.g. a coding sequence or gene) is linked to a regulatory element in such a way that the regulatory element sequence controls and regulates the transcription or translation or both of that polynucleotide sequence. The orientation of the regulatory element may vary (eg, be in reverse orientation with respect to the right ITR). The term also includes having an appropriate start signal (e.g., ATG) in front of the polynucleotide sequence to be expressed and maintaining the correct reading frame to permit expression of the polynucleotide sequence under the control of the expression control sequence and

production of the desired polypeptide or protein. Regulatory sequences can also include 3' sequences which ensure correct termination (eg. polyadenylation stop signal).

5 The term "gene therapy" used herein, refers to a method which introduces a segment of exogenous nucleic acid into a cell in such a way that it results in functional modification to the recipient cell by expression of the exogenous nucleic acid. The exogenous nucleic acid is typically therapeutic in that the expression of the encoded protein, polypeptide or RNA corrects cellular dysfunction due to a genetic error or more generally counteracts any undesirable functions which are associated with a genetic or acquired disease. The  
10 term "exogenous nucleic acid" refers to DNA or RNA sequences not normally expressed in the treated transformed cell. The term also refers to DNA and RNA sequences which are expressed in a treated transformed cell at a higher, lower or in an otherwise different pattern than in the untreated, nontransformed cell. This non-natural expression can also be termed heterologous expression.

15 A "gene therapy vector" refers to a vector used for gene therapy. i.e. to introduce the exogenous nucleic acid into a recipient or host cell. The exogenous nucleic acid may be transiently expressed or integrated and stably expressed in the recipient or host cell.

20 The term "plasmid" as used herein refers to any nucleic acid molecule which replicates independently of the host, maintains a high copy number, and which can be used as a cloning tool.

The term "parallel strand of DNA" and "anti-parallel strand of DNA" refers to as each of  
25 the strands of DNA of the double stranded adenovirus. The Figures diagram the location of certain nucleotides on the parallel strand of DNA. The anti-parallel strand of DNA refers to the other of the two strands of DNA which is not depicted in the Figures. The fiber protein is encoded on the anti-parallel strand of DNA. To simplify the vector diagrams, the fiber sequences are shown on the parallel strand even though the gene is  
30 located on the anti-parallel strand.

The term "reporter gene" refers to any nucleic acid sequence which encodes a polypeptide or protein which can be readily detected by, for example, a visual, microscopic, immunochemical or enzymatic assay. Preferred reporter genes are those  
5 that can be detected by a non-destructive method that does not destroy the treated, transformed cells or tissue.

The term "selection gene" used herein refers to any nucleic acid fragment which encodes a polypeptide or protein whose expression is used to mark a cell as a transformed cell by  
10 a given vector.

The term "therapeutic gene" refers herein to a DNA fragment encoding a functional polypeptide, protein or RNA, which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or to the organ or organism containing the host  
15 cell. The therapeutic benefit may result from modification of a function of a native gene in a host or from the additional function provided by the therapeutic protein, polypeptide or RNA.

The term "host tissue" or "host cell" as used herein, refers to a tissue or cell in which a  
20 therapeutic gene is to be expressed to modify its function.

It is well-known in the biological arts that certain amino acid substitutions may be made in protein sequences without affecting the function of the protein. Generally, conservative amino acid substitutions or substitutions of similar amino acids are tolerated  
25 without affecting protein function. Similar amino acids can be those that are similar in size and/or charge properties, for example, aspartate and glutamate, and isoleucine and valine, are both pairs of similar amino acids. Similarity between amino acid pairs has been assessed in the art in a number of ways. For example, Dayhoff et al. (1978) in *Atlas of Protein Sequence and Structure*, Volume 5, Supplement 3, Chapter 22, pp. 345-352,  
30 which is incorporated by reference herein, provides frequency tables for amino acid substitutions which can be employed as a measure of amino acid similarity. Dayhoff et

al's frequency tables are based on comparisons of amino acid sequences for proteins having the same function from a variety of evolutionarily different sources. Therefore, any obvious changes in the amino acid sequences (as described above) to the sequences of the invention are already contemplated.

5

Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups include but are not limited to: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, asparagine-glutamine, and aspartate-glutamate. Therefore, polypeptide substitution for "substantially similar" sequences (as described above) to the amino acid sequences described invention are already contemplated.

20

In order that the invention herein described may be more fully understood, the following description is set forth.

The present invention provides unique gene transfer vehicles which overcome many of the limitations of prior art vectors. The invention describes a first generation adenovirus vectors comprising left and right Ad ITRs, an Ad packaging sequence, a transgene cassette with regulatory elements, and a pair of cassette ITRs flanking the transgene cassette that direct predictable viral genomic rearrangements during viral replication as well as direct the integration of the transgene cassette into the host cell genome.

30

One predictable rearrangement that occurs during viral replication is the generation of a gutless adenovirus vector (also referred to herein as  $\Delta$ Ad) that comprises right and left Ad ITRs, an Ad packaging sequence, a transgene cassette flanked by cassette ITRs and the gutless vector is devoid of all other immunogenic viral genes.

5

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the transgene cassette is directed by co-infection with an Ad vector expressing e.g., the rep 78 protein to achieve site-specific integration in the e.g., AAVS1 site on human chromosome 19.

10

The invention further describes a novel way of targeting these recombinant adenovirus vectors to selected cells by modifying the adenovirus fiber protein that is expressed on the capsid. Changes to both the fiber shaft and the fiber knob domain proved to successfully retarget the Ad vector to a desired cell type. In addition, the G-H loop within the fiber knob domain is identified as a novel site that affect the binding affinity and specificity of the recombinant adenovirus vector. Substitution of peptide sequences into the G-H loop retarget the gutless vector to a desired cell type.

An adenovirus display library has been generated that expresses random peptides within the G-H loop of the fiber protein. This type of a library is used as ligands to screen for adenovirus vectors that bind to desired cell types. One advantage of using an adenovirus display library versus a phage display library is that once adenovirus affinity to a desired cell is identified the targeted adenovirus vector is ready to accept a transgene cassette and can be used to generate a gutless adenovirus vector, for example for use in gene therapy.

25

The chimeric Ad vectors described below contain a modified fiber protein in the capsid of the adenovirus which renders the vector capable of infecting a desired cell types. Therefore, according to the invention, a gutless chimeric  $\Delta$ Ad-AAV vector can be generated to introduce any transgene(s) into any host cell or tissue which is normally refractory to most commonly used gene therapy viral vectors. In addition, the chimeric  $\Delta$ Ad-AAV vector of the invention, is devoid of adenoviral genes, and contains AAV ITR

30

sequences that flank the transgene cassette, which direct stable transgene integration in the host genome allowing long term expression of the transgene.

The transgene cassette described in the invention may carry a transgene which is either a reporter gene, a selectable gene for in vitro or in vivo selection of transduced cells, or a therapeutic gene. In one embodiment of the invention the reporter transgene can be but is not limited to,  $\beta$ galactosidase. Many reporter genes are commonly used in the art, of which any could be carried as a transgene in the Ad.AAV vector of the invention. Other examples of reporter are genes are GFP and alkaline phosphatase.

The following describes an embodiment of the first generation Ad vectors of the invention having a wild-type capsid and a transgene cassette flanked by cassette ITR sequences; (b) fiber protein that is modified to retarget Ad vectors; and (c) the combination of both technologies that enables the production of chimeric  $\Delta$ Ad vector including a modified fiber protein expressed on the capsid which retargets the base vector to a desired cell type for infection and transgene integration.

A. Integrating Ad hybrid vectors of the invention:

It has been shown that inverted repeats (IRs) inserted into the E1 region of AdE1- vectors can mediate predictable genomic rearrangements resulting in a gutless vector genome devoid of all viral genes. A specific embodiment of such IR-mediated rearrangements is the Adeno-AAV, first generation hybrid adenovirus vector containing AAV inverted terminal repeats (ITR) flanking a transgene cassette. The AAV ITRs mediate the formation of a genome similar to that of the  $\Delta$ Ad.IR genome (Steinwaerder et al., 2000 Journal of Virology).  $\Delta$ Ad vectors devoid of all viral genes stably integrate and transduce cultured cells with efficiencies comparable to e.g. rAAV vectors. The Examples demonstrate by Southern blot analysis that the  $\Delta$ Ad vectors integrate randomly into the host genome.

The Ad vectors of the invention comprise a left Ad ITR, an adenovirus packaging sequence located 3' to the Ad ITR; a transgene cassette located 3' to the packaging

sequence comprising a polyadenylation signal, a transgene, and a heterologous promoter, and flanked by a pair of cassette ITRs. Adenoviral genes used for replication such as E1, E2, E3, E4 are located 3' to the right cassette ITR and a right Ad ITR is located 3' to the replication genes. The vectors of the invention are particularly suited to treat: genetic disorders, cancers, and infectious diseases (such as HIV, emboli, or malaria). Treatable genetic diseases such as hemophilia A and B; cystic fibrosis; muscular dystrophy, and  $\alpha_1$ -antitrypsin disorder are ideal candidates for genetic disease that can be treated by vectors of the invention. A specific example of a therapeutic gene to combat a genetic disorder is gamma-globin to ameliorate sickle cell anemia.

To aid in the selection of transduced cells and characterize the integration site of the transgene cassette, an embodiment of the invention includes the addition of a sequence comprising a bacterial origin of replication, plus a selectable gene. An embodiment of this is an S<sub>Nori</sub> sequence added to the transgene cassette. This allows the  $\Delta$ Ad to be expressed in human and bacterial cells, therefore allowing selection of the transduced cells and characterization of the integration site in the genome of transduced mammalian cells.

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein in 293 cells to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration, which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

Integration of the transgene cassette contained in the adenoviral vectors into chromosomes may be associated with silencing (or blocking) of transgene expression. The silencing of transgenes can be overcome by adding insulator elements to the transgene cassette. For example, HS-4 insulator elements derived from the chicken-globin LCR can function in Ad vectors to shield heterologous promoters from adenoviral enhancers. HS-4 insulators or the *Drosophila* Gypsy gene can also be used to prevent silencing transgenes.

Another embodiment of the invention is to split the transgene cassette into two portions of the transgene each carried in a different recombinant adenoviral vector of the invention. Each portion of the same transgene has an overlapping region of homology. After infection with both vectors, each carrying the different but overlapping portion of the same transgene, homologous recombination event occurs resulting in the reconstitution of the complete transgene which is then expressed. This technique is used to produce hybrid adenoviral vectors that accommodate large inserts including, but not limited to a 13kb genomic hAAT gene or a 12kb  $\gamma$ globin LCR $\gamma$ globin expression cassettes for ameliorating sickle cell anemia (or correcting  $\gamma$ -globin mutations). The formation of the hybrid  $\Delta$ Ad vector genomes, after recombination between two vectors, is more efficient if the overlapping region of homology within the transgene is longer.

An advantage of the present invention is a method to rapidly isolate pure gutless hybrid adenoviral vectors such as  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV<sup>fx</sup> vectors. To minimize the contamination of  $\Delta$ Ad with first generation vectors (Ad vectors) a strategy is described in Example I H. It is anticipated that these approaches will yield the same titer of  $\Delta$ Ad vectors, however the contamination with full-length genome vectors will be less. This improved isolation of the vectors is extremely important to avoid toxic side effects after in vivo application.

B. Tropism modified adenovirus vectors:



The Ad vectors of the invention can be modified so that they target a host cell of interest. There are more than 50 human Ad serotypes (Appendix I), including variants with different tissue selectivity or tropism. It is accepted in the art that different Ad serotypes bind to different cellular receptors and use different entry mechanisms. Most recombinant  
5 adenovirus vectors use adenovirus serotype 5 as the base vector serotype 5 (Ad5) (Hitt, M.M., et al, 1997, Adv. in Pharmacology 40, 137-205). Ad5 infection is primarily mediated by its fiber protein binding to CAR and secondarily by its penton base protein binding to integrin. Due to the lack of CAR and/or integrin expression on many cell and tissue types, Ad5 mediated gene transfer is inefficient in a number of tissues which are  
10 important targets for gene therapy such as endothelia, smooth muscle, skin epithelia, differentiated airway epithelia, brain tissue, peripheral blood cells, or bone marrow. The following describes Ad5 vectors of the invention having a change in infectivity and tropism as a result of altering the fiber protein sequence.

15 The infectivity of different Ad serotypes is limited to a number of human cell lines. Infectivity studies revealed that Ad5 and Ad3 are particularly suitable for infecting and targeting endothelial or lymphoid cells, whereas Ad9, Ad11 and Ad35 efficiently infected human bone marrow cells. Therefore, the knob domain of the fiber protein of Ad9, Ad11 and Ad35 are excellent candidates for retargeting the Ad5 vector to human bone marrow  
20 cells. Other possible serotypes include Ad7.

In the modified fiber protein of the invention the fiber knob domain of the Ad5 fiber has been replaced with another Ad serotype fiber knob domain. An embodiment of the invention is the modified Ad5/35 fiber protein (a recombinant Ad5 vector expressing a  
25 modified fiber protein comprising of a fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad35). The Ad5/35 chimeric fiber protein shows a broader spectrum of infection to a subset of CD34+ cells, including those with stem cell activity. The Ad5/11 chimeric fiber protein (a recombinant Ad5 vector expressing a modified fiber protein comprising the fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad11)  
30 showed similar tropism.

In addition to the knob domain modifications, the invention describes the added advantage of modifying both the fiber shaft domain and the fiber knob domain to produce a shortened fiber protein. The length of the fiber shaft domain plays a key role in the host receptors used for viral vector entry into the host cell. To show this Ad5, Ad5/9, and Ad5/35 variants were constructed with long (22 $\beta$ -sheets) and short-shafted (7 $\beta$ -sheets)-shafted fibers. These analyses demonstrated that efficient viral infection involving CAR as the primary receptor for Ad5, Ad5/9 requires a long-shafted fiber protein, whereas the cell entry strategy of Ad5/35 (which binds to an still uncharacterized non-CAR receptor) does not depend on the shaft length (Fig. 3). The modification in both the fiber shaft domain length (between 5 > 10  $\beta$ -sheets) and the fiber knob domain (from a different Adserotype than the base vector is a novel mode of altering Ad vector tropism.

To broaden the repertoire of cell types that Ad vectors can infect, a specific binding region, the G-H loop, within the knob domain has been newly identified herein to improve binding affinity and specificity. Alteration within this region will redirect the Ad vector to a desired cell type. For example, the invention describes the G-H loop sequence within the fiber protein knob domain, which can be replaced with heterologous peptide ligand sequences without affecting the functionally important tertiary structure of the Ad fiber knob domain, while changing the binding affinity and specificity of the vector (Figs. 6,7). This G-H loop region is exposed on the central part of the knob surface and may be strategically a better site for incorporation of heterologous ligands than the peripheral H-I loop (Krasnykh, V. et al., 1998, *J. Virol.*, 72:1844-52.) of the knob C-terminus (Michael, S. I., et al., 1995, *Gene Ther.*, 2:660-8., Wickham, T. J. et al., 1996, *Nat. Biotechnol.*, 14:1570-3.), which are the substitution sites used by others. Therefore, these G-H loop modifications within the fiber knob domain will allow the Ad vector to be redirected to infect a desired cell type, as long as the G-H loop ligand sequence binds to at least one surface protein on the desired cell type. Fig. 7 shows some possible substitutions. Example II J demonstrates that the virion tolerates the insertion of a cycling peptide (12 amino acids) with a constrained secondary structure that allows the exposure on the knob surface. A defined ligand (RGD) can be inserted into the G-H and the H-I loop of an

Ad5 capsid that is ablated for CAR, and integrin tropism. Infectivity studies show the potential advantage of this new insertion site.

Use of the Vectors of the Invention for "Gene Therapy"

5

The liver is the major organ for protein synthesis. Therefore an important goal of gene therapy is to target gene therapy vectors to the liver. To genetically correct many types of mutant proteins hepatocytes need to be infected with gene therapy vectors carrying a corrected transgene. Example II J describes a G-H loop substitution in the knob domain  
10 of the fiber protein with both R1 and RII+ (of the malaria circumsporozoite surface protein) in a short shafted fiber protein which directs the vector to have affinity and specificity to hepatocytes.

Example II K applies a similar protocol to alter the fiber knob domain in the G-H loop  
15 region with peptides that target the vector breast cancer cell lines (MDA-MB-435). These novel approaches to redirect vectors described in the invention allow lower doses of the gene therapy vectors to be administered with a higher safety profile.

In example II L a protocol for preparing an adenovirus display library is described that  
20 uses the fiber knob protein to display a library of random peptide sequences within the G-H loop. This library of adenoviruses with modified fiber proteins is screened for affinity and specificity for a desired cell type. There are two main advantages of using this adenovirus display library to screen for target peptides that allow binding to a desired cell type over a phage display library system. First, once a ligand peptide is identified that  
25 binds to the desired cell type it is already in the vector of choice for gene therapy delivery. The peptide does not need to be engineered into another vector, as is the case for the phage display library vectors. This reduces the steps required to identify a targeted fiber protein for a desired cell type. The second advantage of this method, is that the adenoviruses are able to display multiple copies of the modified fiber protein on their  
30 capsid. This allows for dimerization and trimerization of the fiber protein with the host cell receptor. The multimerization of fibers proteins is a realistic, in vivo interaction of

the trimeric fiber protein with the host cell receptor. In contrast, phage vectors can only display one fiber peptide sequence on their surface, which significantly limits the ability of interaction with host cell surface receptors.

5     C. A chimeric adenovirus vector with selective tropism:

The chimeric vectors of the invention combine two vectors: an Ad.ITR and a Ad. fx where fx describes a modified fiber protein. A first generation adenovirus vector of serotype 5 is the base vector that carries a transgene cassette flanked by heterologous  
10     ITRs. These specific inverted terminal repeat sequences, such as AAV ITRs direct stable integration of the transgene cassette into the host genome as well as control predictable genomic rearrangements that occur during viral replication. This vector can also carry a modified fiber gene (described in Examples II). During replication predictable genomic rearrangements occur which result in the generation of a gutless adenovirus vector (e.g.  
15      $\Delta$ Ad.AAV<sup>fx</sup>) which expresses the modified fiber protein on its capsid. The modified fiber protein allows the gutless vector to be targeted to a selected cell type. The targeted vector is a gutless adenovirus vector devoid of adenoviral genes which can integrate its transgene into the host genome. The transgene cassette can carry reporter, selectable, or therapeutic genes.

20

In one embodiment of the invention, the gutless targeted  $\Delta$ Ad.AAV<sup>fx</sup> carries the reporter gene of  $\beta$ galactosidase,  $\Delta$ Ad.AAV<sup>fx</sup> -BG). For easy in vitro selection of human and bacterial cells that are transduced with the hybrid Ad vector, a bacterial sequence for the origin of replication can be added to the hybrid Ad vectors. An example of this is  
25      $\Delta$ Ad.AAV<sup>fx</sup> -Snori, in which a SNori sequence is added into the transgene cassette. This site allows for G418 selection on cells infected with  $\Delta$ Ad.AAV<sup>fx</sup> -SNori. This in vitro selection provides a tool to analyze the site of transgene integration and the flanking chromosomal regions. Fluorescent in situ hybridization (FISH) is an alternative method to confirm vector integration.

30

An advantage of the  $\Delta$ Ad chimeric vector for gene transfer is the efficient and stable integration of a large transgene cassette up to e.g. 22kb which is significantly larger than the capacity of retroviral vectors . This is of particular interest for gene therapy. For example, to ameliorate sickle cell anemia  $\Delta$ Ad.AAV<sup>fx</sup>  $\gamma$ globin , an expression transgene cassette with the gamma-globin gene that targets and integrates, can be inserted into bone marrow stem cells for long term expression of the gamma-globin gene.

To achieve site-specific gene integration, rep78 protein is used for transgene integration into the AAVS1 site (described in Example 1 D). However, this may silence transgene expression. To prevent the integrated transgene from being silenced by host genomic elements (such as positional effects or downstream enhancers), LCRs or insulator elements are incorporated into the transgene cassette.

The following examples are presented to illustrate the present invention and to assist one of ordinary skill in making and using the same. The examples are not intended in any way to otherwise limit the scope of the invention.

## **EXAMPLE I**

### **NOVEL ADENOVIRAL VECTOR Ad.AAV**

#### **A. Integrating $\Delta$ Ad.AAV hybrid vectors devoid of all adenoviral genes.**

*In vitro* and *in vivo* studies with rAAV indicate that the only requirement for rAAV integration are the AAV ITRs and as yet unknown host cellular factors. It is thought that specific sequences or secondary structures present in AAV ITRs are prone to integration into host chromosomal DNA. In order to combine advantages of adenoviral vectors (high titer, high infectivity, large capacity) and the integration capability of AAV ITRs, AAV vector DNA with AAV ITRs flanking cassettes a secreted human placental alkaline phosphatase (SEAP) - neomycin phosphotransferase (neo) reporter gene cassette

(Alexander, I.E., et al. 1996, *Gene Therapy*, 7, 841-850) is incorporated into the E1-region of E1/E3 deleted adenoviral vectors (Ad.AAV1) (Figure 1, top).

## METHODS

5

### Production/Characterization of Viral Vectors

#### Plasmids:

- 10 The AAV1 vector cassette containing AAV ITRs and SEAP/neo expression units is obtained by AseI/ScaI digestion of the plasmid pALSAPSN (Alexander, I. E. et al. 1996, *Human Gene Therapy*. 7:841-50). The 4.4kb AAV vector fragment was cloned via NotI adapter linkers into pXJCL1 (Mirobix, Toronto, Canada) (pAd.AAV1). Another shuttle vector (pAd.AAV1- $\Delta$ 2ITRs) lacking the AAV ITRs is generated by inserting the 3.7kb
- 15 AflII/BsmI fragment of pALAPSN into pXJCL1. For pAd.AAV1 $\Delta$ 1ITR, a construct is used where a spontaneous deletion in the left AAV ITRs between the A and A' regions has occurred. To create a second hybrid vector (Ad.AAV2), the AAVSNori cassette developed by E. Rutledge is used. AAV vector DNA obtained is from pASNori (Rutledge, E. A., Russell, D.W. 1997. *Journal of Virology* 71:8429-8436) as a 3.4kb
- 20 BsaI/ScaI fragment and inserted into the EcoRV site of pXCJL1. As it is generally known for AAV vector plasmids, the AAV ITRs are prone to rearrangements. To minimize deletions in these functional critical regions, all constructs for generation of hybrid vectors are assembled in low copy-number plasmids which are grown in E. coli Top10, JC811, or XL1 Bluecells (Stragene, La Jolla, Calif.). Furthermore, after each
- 25 cloning step or large-scale plasmid amplification, both AAV ITRs are carefully mapped by restriction analysis with enzymes that cut inside or adjacent to the ITRs (BssHII, AhdI, SmaI, BglI, BsmI, AflII, and ScaI).

Adenoviruses:

First-generation viruses with the different transgene cassettes incorporated into the E1 region are generated by recombination of the pΔE1aSpl<sup>a</sup>- or pXCJL1-derived shuttle plasmids and pJM17 (Microbix) in 293 cells as described earlier (Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791). For each virus, at least 20 plaques are picked, amplified, and analyzed by restriction digest. Viruses containing two AAV ITDRs tend to rearrange within the ITRs, with other adenoviral sequences, or with adenoviral sequences present in the 293 cell genome. Only plaques from viruses with intact ITRs are amplified, CsCl banded, and titered as described earlier (Kay, M. A., et al. 1995. *Hepatology* 21:815-819; Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960). All virus preparations tested are negative for RCA and bacterial endotoxin (Lieber, A., et al. 1997. *Journal of Virology* 71:8798-8807). Virus is stored at -80°C in 10mM Tris-Cl, pH 7.5-1 mM MgCl<sub>2</sub>-10% glycerol.

To generate ΔAd.AAV, 293 cells are infected with Ad.AAV1 at an multiplicity of infection (MOI) of 25 and harvested 40h after infection. Cells are lysed in PBS by 4 cycles of freeze/thawing. Lysates are centrifuged to remove cell debris and digested for 30 min at 37°C with 500 units/ml DNaseI and 200 μg/ml RNaseA in the presence of 10mM MgCl<sub>2</sub>. 5ml of lysate is layered on a CsCl step gradient (0.5ml - 1.5g/cm<sup>3</sup>, 2.5ml - 1.35g/cm<sup>3</sup>, 4ml - 1.25g/cm<sup>3</sup>) and ultracentrifuged for 2h at 35,000 rpm (rotor SW41). CsCl fractions are collected by puncturing the tube and are analyzed for viral DNA (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960; Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) or subjected to ultracentrifugation at 35,000 rpm for 18 hours in an equilibrium gradient with 1.32 g/cm<sup>3</sup> CsCl. The band containing the deleted viruses ΔAd.AAV is clearly separated (0.5cm distance) from other banded viral particles containing full-length ad.AAV genomes. ΔAd.AAV1 fractions are dialyzed against 10mM Tris-Cl, pH 7.5-1mM MgCl<sub>2</sub>-10% glycerol and stored at -80°C. The genome titer of ΔAd.AAV1 preparations is determined based on quantitative Southern analysis of viral DNA purified from viral particles in comparison to different concentrations of a 4.4kb AseI/ScaI fragment of pALSAPSN according to a protocol described earlier (Lieber, A.,

et al., 1996, *J. of Virology*, 70, 8782-8791). In total, the production of  $1 \times 10^{13}$  genome-particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of actual work.

Titers routinely obtained are in the range of  $3-8 \times 10^{12}$  genomes per ml. Assuming one genome is packaged per capsid, the genome titer equals the particle titer. The level of contaminating Ad.AAV1 is less than 0.1% as determined by Southern analysis, which is consistent with results obtained by plaque assay on 293 cells (fewer than 5 plaques per  $10^6$  total genomes). The primers used for sequencing the left and right ITR-vector-junction are

5'GGCGTTACTTAAGCTAGAGCTTATCTG, and  
5'CTCTCTAGTTCTAGCCTCGATCTCAC.

The recombinant AAV virus stock containing the SEAP/neo cassette (AV2/ALSAPSN, [Alexander, I. E. et al, 1996, *Human Gene Therapy*, 7:841-50] used in these studies were obtained from Dusty Miller (FHCRC, Seattle). The stock was free of contaminating replication competent AAV (<50 particles/ml) and wildtype adenovirus (<100 particles/ml). The genome titer of the virus stock was obtained by quantitative Southern Blots as described by Russell et al. (Russell, D. et al. 1994 *Proc. Natl.Acad.Sci. USA* 91:8915-8919).

#### Electron Microscopy:

For examination of viral particles in the transmission electron microscopy studies, CsCl-purified virions are fixed with glutaraldehyde and stained with uranyl acetate as described previously (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960).

#### *RESULTS*

During replication of these hybrid vectors in 293 cells, a 5.5kb genome ( $\Delta$ Ad.AAV1) is efficiently generated and packaged into adenovirus (Ad5) capsids. The  $\Delta$ Ad.AAV1



genome contains the left adenovirus ITR and the packaging signal followed by the AAV-vector cassette and a duplicate of the adenoviral packaging signal and ITR in reverse orientation (Figure 1, bottom). The hybrid vector is devoid of all viral genes, thus eliminating toxic effects and the elicitation of cellular immune responses. The spontaneous formation of the small hybrid vector genome  $\Delta$ Ad.AAV1 requires the presence of two intact AAV ITRs and does not occur with partly deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette.

Hybrid vectors containing different transgenes:

To construct a hybrid vector with a transgene that can be detected in situ the SEAP/neo expression unit in Ad.AAV1 is replaced by the *E. coli*  $\beta$ -galactosidase gene. This hybrid vector is named  $\Delta$ Ad.AAV1. During generation of the corresponding plasmid constructs the AAV ITR sequences tend to rearrange and abolish their functional properties. This problem can be circumvented by using low copy number plasmids as cloning vectors grown in bacteria strains depleted for all recombination proteins (e.g. JC811). Furthermore, the intactness of both AAV ITRs after each cloning step can be examined for characteristic endonuclease digestion. Recently, another hybrid vector  $\Delta$ Ad.AAV1Nori has been generated which contains the neo gene under the control of both the simian virus 40 (SV40) early promoter and the transposon 5 (Tn5) promoter for expression in human and bacterial cells, as well as the p15A bacterial replication origin with the direction of the leading strand DNA synthesis opposite that of neo gene transcription. Thus, SNori can be used for G418 selection of integrated vector in eukaryotic cells as well as for rescue of vector together with flanking host DNA after integration. The recovered plasmids can be propagated in *E. coli* under selection with kanamycin due to the bacterial origin and the neo gene. SNori containing vectors allow a rapid estimation of total integration events based on the number of G418 resistant colonies. Moreover, vector DNA together with flanking chromosomal DNA can be rescued as plasmids from single G418 resistant clones and can be used for sequencing to determine integration junctions. Both hybrid vectors are produced at a titer of about

$3 \times 10^{12}$  genomes per ml. The ratio of genome titer to transducing particles for  $\Delta$ Ad.AAVBG is  $\sim 200:1$  based on  $\beta$ -Gal expression.

## DISCUSSION

5

$\Delta$ Ad.AAV1 could spontaneously form during adenovirus replication. Another possible mechanism of  $\Delta$ Ad.AAV1 formation is based on the unique mechanism by which adenovirus replicates its genome (van der Vliet, B., 1995, *In* w. Doerfler, et al. (eds.) vol. 2 p. 1-31, Springer-Verlag, Berlin) (see Figure 1). Ad DNA replication is initiated by the TP/pTP (terminal protein) that binds to specific sites within the ITRs on both ends of the linear genome and catalyzes, in complex with Ad pol, the binding of the 5' CTP, the first nucleotide of the daughter strand. DNA synthesis proceeds in a continuous fashion to the other end of the genome (Figure 1A). Only one of the DNA strands serves as template. One of the replication products is a single-stranded DNA that circularizes through annealing of its self-complementary ITRs. The resulting duplex "panhandle" has the same structure as the termini of the duplex viral genome that allows the binding of pTP and the initiation for synthesis of a complementary strand using the single-stranded "panhandle" molecule as template (Figure 1C). In the case of Ad.AAV1, the Ad pol synthesizes the single strand of the adenoviral genome starting from the left Ad ITR until it reaches the second AAV ITR. During synthesis of the second AAV ITR a certain percentage of the single-stranded molecules form a loop hybridizing to the complementary region within the first AAV ITR that was replicated earlier, allowing Ad pol to use the same viral DNA strand to read back towards the left ITR (Figure 1B). The resulting "panhandle" structure can be resolved in a similar way as a full-length intermediate shown in Figure 1C, generating a double stranded, linear molecule with the above described structure that can be packaged into Ad virions. The ratio of viral DNA to protein concentration in purified  $\Delta$ Ad.AAV1 particles is comparable to that obtained from Ad.AAV1 particles. This indicates that despite the smaller size, only one  $\Delta$ Ad.AAV1 genome is packaged, resulting in particles with a lighter buoyant density ( $\sim 1.32 \text{ g/cm}^3$ ). Electron microscopy demonstrates the icosahedral shape of  $\Delta$ Ad.AAV1 particles (Figure 2). Staining with uranyl acetate causes the central viral cores to appear

30

electron dense.  $\Delta$ Ad.AAV1 virions have only a spotted luminal dark staining as expected with only one 5.5kb genome being packaged per capsid.

**B. In vitro  $\Delta$ Ad.AAV1 production:**

**Characteristics of deleted adeno-AAV vectors ( $\Delta$ Ad.AAV):**

A number of experiments to clarify the mechanisms of  $\Delta$ Ad.AAV genome formation are carried out. Specifically, the presence of two intact AAV ITRs flanking a reporter gene cassette is required for the effective formation of  $\Delta$ Ad.AAV genomes. This process does not occur with partially deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette. Furthermore, in vitro transduction studies are performed with different genome titers of  $\Delta$ Ad.AAV1, Ad.AAV1, and Ad.AAV1- $\Delta$ 2ITRs, (lacking the two AAV ITRs) which determine the number of G418 resistant colonies that formed after 4 weeks of selection (Table I).

$\Delta$ Ad.AAV1 is routinely produced at a high titer ( $5 \times 10^{12}$  genomes per ml with  $>10^4$  produced genomes per 293 cell) and at a high purity with less than 0.1% contaminating full length Ad.AAV1 genomes by a technique normally used for amplification and purification of recombinant adenovirus.

**In vitro transduction studies with hybrid vectors on CD34+ cells and erythroleukemia cells:**

In order to test whether the hybrid vectors allow for gene transfer into cell types, that have to be targeted for sickle cell therapy, infection/transduction studies are performed using CD34+ enriched human bone marrow cells, derived from mobilized peripheral blood and the human erythroleukemia cell line K562 which express  $\epsilon$  and  $\gamma$  globin genes.

**METHODS**

Cell Culture:

SKHep1 cells (HTB-52, American Type Culture Collection, Rockville, MD), an endothelial cell line derived from human liver [Heffelfinger, S.C., et al., 1992, *In vitro* 5 *Cell Dev. Biol.* 28A, 136-4-142], are grown in high-glucose Dulbecco's modified Eagle medium with 10% fetal calf serum. SKHep1 cells are analyzed for integrated AAV provirus by Southern analysis of genomic DNA using the AAV1 wild type genome obtained from pAAV/Ad (Samulski, R. J., et al. 1989. *Journal of Virology* 63:3822-3928) (gift from David Russell, University of Washington) as a probe. No specific bands are 10 detected in undigested genomic SKHep1 DNA or after digestion with HindIII. For viral infection, confluent cells are incubated with different viral doses for 2 hours, followed by intensive washing. For G418 selection, 24h after infection with  $\Delta$ Ad.AAV1, SKHep1 cells are trypsinized and plated at different dilutions under G418 selection (900 $\mu$ g/ml active compound, Boehringer-Mannheim, Germany). G418 containing culture medium is 15 changed every 3 days. The number of colonies with  $>10^4$  cells is counted after 4 weeks of selection and divided by the number of initially seeded cells. This ratio is used to express the integration frequency of  $\Delta$ Ad.AAV1. Single colonies are obtained by limiting dilutions of infected cells in 96 well plates. Colonies are expanded to  $1 \times 10^6$  cells in the presence of G418. Immunofluorescence analysis for adenoviral proteins expressed in 20 SKHep1 cells 3 days post-infection is performed as described earlier [Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791].

RESULTS

25 293 cells are infected with the first generation vector Ad.AAV1. During replication of Ad.AAV1, the small  $\Delta$ Ad.AAV1 genome forms spontaneously and is packaged into adenovirus capsid. At 36 hours after infection cells are harvested and virus is released by several cycles of freeze/thawing. The mixture of Ad.AAV1 and  $\Delta$ Ad.AAV1 particles in the cell lysate is then separated by ultracentrifugation in a CsCl step gradient. Due to its 30 lighter buoyant density, the band containing the  $\Delta$ Ad.AAV1 particles is clearly separated (0.8cm distance) from the band containing full-length virus (Lieber, A., et al. 1999. *J*

*Virol* 73:9314-24).  $\Delta$ Ad.AAV1 is purified further by an additional CsCl equilibrium gradient and is stored in 10mM Tris pH7.5, 10% glycerol, 1mM MgCl<sub>2</sub> in 80°C. In total, the production of  $2 \times 10^{13}$  (genome) particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of work. All functions for  $\Delta$ Ad.AAV1 replication and particle formation are provided from

5 Ad.AAV1 genomes amplified in the same cell. The efficiency of vector production measured on a genome-per-cell-basis is comparable or higher than labor-intensive, newer techniques for rAAV production, which have not yet been proven to be reliable. The estimated ratio of transducing/genome titer for  $\Delta$ Ad.AAV1 is 1:200 (based on SEAP expression at day 3 post-infection), whereas for the average rAAV preparation, it is in the

10 range of  $1:10^3$  to  $1:10^4$ .  $1 \times 10^5$  confluent SKHepl cells are infected with different MOIs of rAAV1 (stock:  $1 \times 10^{10}$  genomes per ml),  $\Delta$ Ad.AAV1 (stock:  $5 \times 10^{12}$  genomes per ml), Ad.ADAV1 (stock:  $1 \times 10^{13}$  genomes per ml), and AdAAV1.2ITR (stock:  $9 \times 10^{12}$  genomes per ml), in a volume of 100ml 24 hours after infection, cells are washed, trypsinized, and plated at different dilutions. G418 is added 24 hours after plating and selection is

15 performed for 4 weeks. G418 resistant colonies contain on average  $>5 \times 10^4$  cells (at least 16 cell divisions). A significant number of small colonies visible at 2 weeks post-infection do not survive continued selection, probably due to episomal vector expression. Cells infected with first-generation adenoviruses with MOIs greater than  $1 \times 10^4$  develop CPE during the first week of selection. The rAAV titer is not high enough to perform

20 infection studies with MOIs greater than  $10^4$ . The colony formation is expressed as percentage of the number of colonies after selection to the number of cells initially seeded for selection (Table I).

**TABLE I**

25

**Formation of G418 resistant colonies after infection with hybrid viruses in comparison with rAAV.**

| MOI<br>(genomes per<br>cell) | Formation of G418 resistant colonies in % (SEM)<br>(after 4 weeks of selection) |
|------------------------------|---|
|------------------------------|---|

|                 | rAAV1      | Ad.AAV1    | Ad.AAV1    | Ad.AAV1<br>2ITRs |
|-----------------|------------|------------|------------|------------------|
| 10 <sup>1</sup> | 0          | 0          | 0          | 0                |
| 10 <sup>2</sup> | 0          | 0          | 0          | 0                |
| 10 <sup>3</sup> | 2.7 (1.6)  | 1.3 (1)    | 5.4 (3.0)  | 0                |
| 10 <sup>4</sup> | 90.8 (7.0) | 48.0 (8.9) | 12.9 (7.2) | 0                |
| 10 <sup>5</sup> | N/A        | 93.1 (5.4) | 3.8 (2.1)  | 0                |
| 10 <sup>6</sup> | N/A        | 100        | 0          | 0                |
| 10 <sup>7</sup> | N/A        | 100        | 0          | 0                |

N = 3 (SEM is indicated in parentheses.)

K562 cells are infected with different MOIs of  $\Delta$ Ad.AAVBG (1-10<sup>8</sup> genomes per cell):

- 5 Three days after infection, the total number of viable cells (based on Trypan blue staining) and the percentage of infected cells (based on X-Gal staining) are determined for all MOIs. The results are presented in Figure 4A.

Initial integration studies:

10

K562 cells are incubated with  $\Delta$ Ad.AAVSNori at an MOI of  $2 \times 10^5$  genomes per cell and the colonies that formed after 4 weeks of G418 selection are counted in 96 well plates. G418 resistant colonies contain on average  $>5 \times 10^4$  cells which means that the original cell underwent at least 16 cell divisions.

15

Infection studies with Ad.AAVBG (1-10<sup>8</sup> genomes per cell) on CD34+ cells:

- 20 Cell infection on CD34+ are as described for K562 cells. CD34+ cells are cultured in IMDM supplemented with 20% FCS, kit ligand (stem cell factor-SCF) (100ng/ml), and IL-3 (100ng/ml). Since a number of reports suggest that specific cytokines like GM-CSF or M-CSF which induce stem cell differentiation can stimulate integrin expression and may therefore affect internalization of Ad5 vectors, infection rates are compared with

Ad5 based hybrid vectors on CD34+ cells cultured with and without pre-stimulation with GM-CSF (50ng/ml) or M-CSF (50U/ml). The number of infected cells is counted based on X-Gal staining at day 3 after infection. To test for dose-dependent toxicity, viable cells are counted (based on trypan blue exclusion) at day 3 post-infection. Furthermore, whether high viral doses affect the ability of CD34+ cells to differentiate in methyl cellulose colony assays in presence of IL-3 and SCF is analyzed. The results are expressed as viable cells/X-Gal positive cells vs MOI (see Figure 4B).

## DISCUSSION

The above data demonstrate that  $\Delta$ Ad.AAV transduces stably an immortalized human cell line with a low frequency comparable to rAAV, however, transduction rates could be scaled up to 100% by using greater MOIs of  $\Delta$ Ad.AAV1, which is produced at higher titers than rAAV1. In contrast to infection with the first-generation vector, Ad.AAV1, infection with  $\Delta$ Ad.AAV1 is not associated with dose-dependent cytotoxicity because no viral proteins are expressed from these vectors in transduced cells. Furthermore, viral proteins present in the incoming  $\Delta$ Ad.AAV1 particles are not problematic in the dose range use. The comparison of transduction rates of  $\Delta$ Ad.AAV/Ad.AAV1 with the vector lacking AAV ITRs, Ad.AAV- $\Delta$ 2ITRs, supports the hypothesis that the presence of two intact AAV ITRs is crucial for hybrid vector integration.

The data demonstrate that the leukemia cell line can be infected at ~90% efficiency with Ad5 based hybrid vectors at MOIs  $2 \times 10^5$  genomes per cell without significant toxic side effects. However, this dose is still ~100 times greater than the dose necessary to infect 100% of HeLa cell, hepatoma cells, primary hepatocytes and other cell lines generally considered as permissive for Ad5 vector infection.

Since viral DNA in cells infected with  $2 \times 10^4$  genomes (or 100 transducing particles per cell) should be lost after 7 cell divisions, the presence of G418 resistant cells in the observed colonies suggests that  $\Delta$ Ad.AAVSNori genomes are integrated into or stably associated with the host genome. Based on the number of G418 resistant colonies one

out of 25,000  $\Delta$ Ad.AAVSNori genomes integrates stably into K562 cells. This is in agreement with the results obtained earlier with  $\Delta$ Ad.AAV1 in SK Hep1 cells.

5 The maximal dose used for infection of CD34+ cells ( $1 \times 10^8$ ) results in X-Gal staining of only ~10% of cells independently of GM-CSF/M-CSF. This demonstrates the obvious inability of Ad5 to infect CD34+ cells and is probably caused by the absence of specific receptors and/or integrins on the cell surface. CD34+ cells tolerate a large range of viral doses ( $1-10^7$ ) without obvious effects on cell viability and total cell number. This is not surprising because in order to develop toxic side effects adenovirus has to enter the cell and express viral genes. Hybrid vectors can be produced at titers of  $5 \times 10^{12}$  genomes per ml. Thus, the maximal MOI that can be used for infection (of  $10^4$  cells) is  $\sim 5 \times 10^8$  (in 100 $\mu$ l storage buffer). Based on the infection studies with  $\Delta$ Ad.AAVBG this dose may not be sufficient to efficiently transduce CD34+ cells and to obtain an appreciable number of G418 resistant colonies.

15

### C. In Vivo Properties Of $\Delta$ Ad.AAV1:

Viral DNA is labeled with BrdU during virus amplification to investigate cellular/nuclear vector uptake in situ. For transduction studies, confluent SKHep1 cells (a human endothelial cell line) are infected with 2000 genomes  $\Delta$ Ad.AAV1 or Ad.AAV1 per cell. BrdU tagged viral DNA is detected in 100% of nuclei at 3 hours post-infection for both viruses indicating efficient cellular and nuclear uptake of hybrid virus DNA.

### *RESULTS*

25

The  $\Delta$ Ad.AAV1 vector transduces a cell in vitro forming G418 resistant colonies with an efficiency of 17 or 58%, after infection with an MOI of  $1 \times 10^3$  or  $1 \times 10^4$  genomes per cell, respectively. Approximately  $2 \times 10^4$   $\Delta$ Ad.AAV1 genomes are required to yield one stable transfectant. Since all stable colonies contain integrated  $\Delta$ Ad.AAV1 vector DNA, this number reflects the minimal integration frequency of  $\Delta$ Ad.AAV1 in SKHep1 cells which is comparable with that from rAAV (Rutledge, E. A. et al., 1997, *Journal of Virology*,

30



71:8429-36). The number of G418 resistant colonies does not necessarily represent the total frequency of integration events because not all integrated copies express neomycin phosphotransferase, due to chromosomal position effects or incomplete integration.

5 The absence of adenoviral gene products in  $\Delta$ Ad.AAV1 transduced cells at day 3 post-infection is demonstrated by immunofluorescence with antibodies to the major late proteins (hexon, fiber) and early proteins (DBP.E4-orf6). Expressed adenoviral proteins are detected only in cells infected with Ad.AAV1. The fact that cells infected with  $\Delta$ Ad.AAV1 do not express potentially cytotoxic adenoviral proteins is important.

10

While an MOI of  $1 \times 10^4$  genomes per cell of the first generation vector Ad.AAV1 induce cytopathic effects in SKHep1 cells at day 3 p.i., no toxic side effects are observed when SKHep1 cells are infected with  $\Delta$ Ad.AAV1 at a dose of up to  $1 \times 10^8$  genomes per cell. Since the transduction efficiency is clearly dose dependent,  $\Delta$ Ad.AAV1 (which can be  
15 produced at titers of  $>5 \times 10^{12}$  genomes/ml) is able to stably transduce Ad5 permissive cell lines or tissues with a 100% efficiency without associated toxicity.

Southern analysis indicates that  $\Delta$ Ad.AAV1 integrates randomly as head-to-tail tandem repeats into the host cell genome via the right AAV ITR, whereas the other junction with  
20 the chromosomal DNA is variable and occurs somewhere within the transgene cassette. In order to confirm the integrated status of  $\Delta$ Ad.AAV1 DNA, high-molecular-weight chromosomal DNA is separated by pulse field gel electrophoresis (PFGE), followed by Southern analysis with a SEAP specific probe (Figure 3). Undigested DNA from control SKHep1 cells give an endogenous SEAP signal that co-migrates with chromosomal DNA  
25 just below the well (lanes 1 and 5). No high-molecular weight episomal forms of  $\Delta$ Ad.AAV1 DNA are detected, whereas a distinct 35 kb band is visible in DNA from SKHep1 cells isolated 3 days after infection with first generation adenovirus, Ad.AAV1 (lanes 4 and 13). Digestion with EcoRI reveals the 4.4kb fragment, which is specific for integrated tandem copies of the AAV cassette (lanes 8 and 12). To eliminate the  
30 possibility that chromosomal DNA is trapped in the well, DNA samples are digested with intron-encoded endonucleases PI-SceI or I-CeuI (Gibco-BRL, Grand Island, NY) with a

sequence specificity or more than 11bp or 9bp respectively. Digestion with PI-SceI yields a >2mb endogenous SEAP signal in SKHep1 cells (lane 2) and an additional signal in the range of ~1mb in G418 resistant colonies transduced with  $\Delta$ Ad.AAV1 (lane 7). I-CeuI digestion results in a smear between 250-1000 kb in  $\Delta$ Ad.AAV1 transduced SKHep1-cells (lanes 10, 11) indicating random integration, whereas a high-molecular weight band specific for the endogenous SEAP gene is observed in control SKHep1 cells (lane 9).

One day after intraportal infusion of  $1 \times 10^{12}$   $\Delta$ Ad.AAV1 genomes in C57Bl/6 mice, BrdU labeled vector genomes can be detected in 85% hepatocytes (Lieber, A., et al. 1999. *J Virol* 73:9314-24). Hepatocellular DNA analysis performed at 2 months post-infusion reveals  $\Delta$ Ad.AAV1 DNA integrated with an average of 0.5 copies per cell into the mouse genome (Lieber, A., et al. 1999. *J Virol* 73:9314-24). To assess potential side effects of intraportal  $\Delta$ Ad.AAV1 infusion, serum glutamic pyruvic transaminase (SGPT), a sensitive marker for hepatocellular injury, is measured for 7 consecutive days post-infusion in combination with histological analysis of liver sections. No significant elevation in SGPT levels, or histological abnormalities are detected after intraportal infusion of  $1 \times 10^{12}$  or  $1 \times 10^{13}$   $\Delta$ Ad.AAV1 genomes, whereas infusion of the same dose of full-length Ad.AAV1 vector is associated with severe hepatotoxicity or fatal outcome. This suggests that the dose of  $\Delta$ Ad.AAV1 administered to mice can be increased to obtain higher transduction efficiencies in vivo without adverse side effects, which is not possible for first generation adenoviruses. Importantly,  $\Delta$ Ad.AAV1 transduced quiescent hepatocytes in vivo, which suggests that integration of hybrid vector DNA may not require cell proliferation. Recently, more detailed in vivo transduction studies with Ad.AAV1 and  $\Delta$ Ad.AAV1 have been performed in Balb/c mice to study whether the absence of adenoviral gene expression in cells infected with  $\Delta$ Ad.AAV1 can avoid an anti-viral immune response and can prolong vector persistence. In this mouse strain, vector DNA is cleared from the liver at 4-6 weeks after infusion with first generation adenoviruses, mostly due to a CTL response against viral proteins produced in transduced cells. Vector DNA is analyzed by genomic Southern Blot of hepatic DNA at 12 weeks

after infusion of  $1 \times 10^{12}$  genomes Ad.AAV1 or  $\Delta$ Ad.AAV1. At this time point, no vector specific signal is detectable in hepatic DNA from mice infused with the first generation vector Ad.AAV1, while  $\sim 0.3$  copies of  $\Delta$ Ad.AAV1 genomes per cell are present in livers of mice that received the hybrid vector, again indicating the superior in vivo properties of the hybrid vector.

#### **D. Effects of Rep Coexpression on $\Delta$ Ad.AAV Integration**

##### **Rep expression after plasmid transfection:**

In order to test whether Rep expression enhances site-specific integration of  $\Delta$ Ad.AAV1 in human cells, a series of Rep expression plasmids are constructed.

#### **METHODS**

The Rep ORF 68/78 (nt 285-2313) including the internal p19 and p40 promoters is obtained from pAAV/Ad (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) by digestion with BsaI/BsrI. This fragment deleted for the AAV p5 promoter is cloned via adapter linkers under RSV or PGK promoter in front of the bovine growth hormone polyadenylation signal (bPA) into pAd.RSV or pAd.PGK (Lieber, A., and Kay, M.A., 1996, *J. of Virology*, 70, 3153-3158; Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) correspondingly.

#### **RESULTS**

The resulting plasmids (pRSVrep, pPGKrep) are transfected into 293 cells or SKHep1 cells, most of the Rep proteins expressed from the heterologous promoters (RSV or PGK) are Rep 68 and Rep 78, while transfection of the rep gene under aP5 promoter (pAAV/Ad) results in predominant Rep 52/40 expression. Thus, transfection of pRSVrep and pPGKrep is more pronounced suggesting a strong transactivation of AAV promoters

by Ela which is produced in 293 cells. This result indicates that minimum expression of rep proteins is necessary to avoid interference with adenovirus replication.

Rep-mediated site-specific integration of  $\Delta$ Ad.AAV1.

5

The potential for site-specific integration is an important characteristic of the novel Ad.AAV vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration, which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

A preliminary test can be performed to confirm the functional activity of Rep 68/78 expressed from pRSVrep to mediate site-specific integration of  $\Delta$ Ad.AAV1 (Figure 5 and 6). Human SKHep1 cells are transfected with pRSVrep or control plasmid (pRSVbGal (Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) (transfection efficiency was ~20%), followed by infection with  $\Delta$ Ad.AAV (2000 genomes per cell). Three days after infection, cells are trypsinized, embedded in agarose, lysed in situ, digested with I-CeuI (an intron-encoded endonuclease with a recognition sequence of more than 10nt), subjected to pulse field gel electrophoresis in 1% agarose gel, and analyzed by Southern Blot. Hybridization with a probe covering the AAVS1 integration site (1.7kb EcoRI/BamHI fragment from the chromosome 19 locus (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) ) reveals an AAVS1-specific band (~240kb) in I-CeuI digested DNA from cells after control plasmid transfection (pCo) +  $\Delta$ Ad.AAV1 infection. An additional signal in the range of 280kb appears in rep expressing cells infected with  $\Delta$ Ad.AAV1

(pRSVrep + VAd.AAV1) indicating a site-specific insertion into the AAVS1 site in a certain percentage of cells. The presence of vector DNA in this 280kb band is confirmed by rehybridization of the same filter with a transgene (SEAP) specific probe. Randomly integrated  $\Delta$ Ad.AAV1 vector appears as a diffuse SEAP signal in the range 280-680kb  
5 (pCo+ $\Delta$ Ad.AAV1, pRSVrep+ $\Delta$ Ad.AAV1). The specific ~1.9 mb band on blots hybridized with the SEAP probe represents an I-CeuI fragment containing the endogenous human SEAP gene.

Incorporation Rep 68/78 function into hybrid vectors to stimulate site-specific integration  
10 Rep overexpression inhibits adenovirus DNA replication, prohibiting the generation of rep expressing Ad vectors using conventional strategies. To solve this problem, significant Rep 68/78 expression from the hybrid vector in virus producer (293) cells must be prevented while maintaining transient Rep expression in target cells (HSC) to mediate site-specific integration. Our hypothesis is that the specific structure of the  
15  $\Delta$ Ad.AAV hybrid virus can be used to bring the rep gene 68/78 into a transcriptionally active position under control of a HSC specific promoter only at late stages of virus replication in 293 cells. This will allow amplification of the hybrid vector in 293 cells, generating high titer virus which activates the incorporated Rep 68/78 functions only in HSC. The general outline of our strategy to produce Rep expressing hybrid vectors is  
20 illustrated in Figure 7. The rep/transgene cassette is assembled based on the left-hand shuttle plasmid used for recombinant adenovirus production. The gene encoding Rep 68/78 is cloned in 3' to 5' orientation in front of a transgene expression cassette flanked by AAV ITRs. Between the transgene cassette and the right AAV ITR an HSC-specific promoter is inserted with direction towards the adenoviral E2, E3, and E4 genes. The  
25 recombinant genome is produced by recombination in E.coli and transfection into 293 cells generates virus (Ad.AAV-rep). The specific structure of  $\Delta$ Ad.AAV with duplicated sequences flanking the AV ITRs is used to bring the rep gene into a transcriptionally active position under control of a HSC specific promoter only during late stages of viral DNA replication in 293 cells. During amplification of Ad.AAV-rep, the smaller genome  
30  $\Delta$ Ad.AAV-rep is formed and packaged into particles, which can be separated by ultracentrifugation in CsCl gradients. The specific structure of  $\Delta$ Ad.AAV-rep brings the

rep gene into 5'→3' orientation in relation to the HSC specific promoter, allowing rep transcription in target cells. After transduction of HSC with purified ΔAd.AAV-rep particles, rep expression is activated and mediates rescue of the AAV-ITR/transgene cassette from the adenoviral vector backbone and site-specific integration. The hypothesis is that Rep-mediated integration into AAVS1 occurs via the right or both AAV/ITRs causing the rep gene to become separated from the hepatocyte-specific promoter once the vector is integrated (Figure 7). Therefore, rep expression should be only transient without critical cytotoxic side effects on the host cell.

10 Promoters that can regulate rep expression:

Potential candidate promoters to drive rep expression with high specificity for HSC and minimal activity in 293 cells are the 454nt CD34 promoter (Krause, D.S., et al., 1997, *Experimental Hematology*, 25, 1051-1061; Yamaguchia, Y. et al., 1997, *Biochimica et Biophysica Acta*, 1350:141-6), the 300nt HS 40 enhancer (Chen, H.L., et al., 1997, *Nucleic Acids Res.* 25, 2917-2922) or a 3kb CD34 enhancer (May, G. et al., 1995, *EMBO J.*, 14:564-74) in combination with an initiator, or the HIV LTR. An optimal promoter is selected based on studies of transient reporter gene expression after plasmid transfection in 293 cells and hepatocytes. All promoters to be tested are cloned in front of the human α<sub>1</sub>-antitrypsin (hAAT)- bovine growth hormone polyadenylation signal (bPA) into the adenoviral shuttle plasmid pCD2 (pAd.-hAAT). Promoter activity can be tested in transient plasmid transfection assays in CD34+ and 293 cells. The promoter with the highest hAAT levels in CD34+ or K562 cells and the lowest hAAT expression in 293 cells is selected for further studies. If high background expression in 293 cells from these promoters is seen, insulators to shield HSC-specific promoters from the Ela enhancer which is still present in Ad shuttle plasmids can be utilized.

Rep genes:

30 The large Rep 68/78 proteins are sufficient to mediate rescue and site-specific integration. Unregulated Rep 52 and Rep 40 expression from the AAV p19 promoter

located within the ORF of Rep 68 and 78 must be prevented because production of these smaller Rep proteins in 293 cells will affect cell viability and adenoviral DNA synthesis. To do this, constructs obtained from Surosky et al., containing a mutated Rep 52/40 start codon to express Rep 68 and 78 individually under CMV promoter can be used. The 293  
5 cells transiently expressing Rep68 or Rep 78 from these constructs can be coinfectd with  $\Delta$ Ad.AAV1 (infection 24 hours after pCMVRep transfection, MOI  $2 \times 10^5$  genomes/cell). Three days after  $\Delta$ Ad.AAV1 infection, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE as described earlier. Efficient Rep mediated excision of the AAV cassette and site-specific integration without flanking adenoviral  
10 sequences are expected and the plasmids pCMVRep68 or pCMVRep78 can be used as a source for the corresponding rep genes and clone them into hybrid vectors.

Vectors:

15 The rep/transgene cassette can be assembled based on pXCJL (Microbix, Toronto). A set of control hybrid vectors can be generated with the AAV-ITR-transgene cassette only without the rep gene. The recombinant Ad.AAV-rep genome can be generated by recombination of the left hand shuttle plasmids with pCD1, a pBHG10 (Microbix, Toronto) derivative, which contains the Ad5 genome deleted for the E1/E3 regions in  
20  $recA^+$  E. coli (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Compared to the standard technique based on plasmid recombination in 293 cells, this approach has the advantage that plaques with recombinant virus appear 3 times faster and the production of illegitimate recombinants is minimized. This allows efficient viral DNA amplification and packaging to occur before Rep expression reaches levels that are potentially  
25 inhibitory for adenoviral replication. The critical variables in maximizing the output of the vector deleted for all adenoviral viral genes are the initial multiplicity of infection and the time of harvesting. These parameters can be optimized for production of  $\Delta$ Ad.AAV-rep hybrid vectors. A number of  $\Delta$ Ad.AAV vectors can be constructed incorporating rep gene. Cryptic promoter and enhancer elements present in the 5'-342nt of the adenoviral  
30 genome can interfere with transgene expression from the heterologous promoters. This is crucial for the strategy to avoid rep expression from  $\Delta$ Ad.AAV-rep genomes in 293 cells.

To ensure efficient transgene expression, insulator fragments such as the chicken beta-globin insulator can be used with a selected promoter, constitutive or inducible.

Rep protein co-packaging:

5

As an alternative to producing hybrid vectors containing the rep 68/78 gene, studies are designed to see whether Rep protein can be co-packaged into  $\Delta$ Ad.AAV capsids and whether these co-packaged Rep molecules are sufficient to mediate rescue and site-specific integration of the AAV-ITR-transgene cassette. Our hypothesis is that the Rep  
10 68/78 binds to the Rep binding site (RBS) present in double-stranded  $\Delta$ Ad.AAV genome and that this complex is co-packaged into adenoviral capsids which are spacious enough to accommodate extra proteins. Based on protein/DNA ratio analysis performed previously in purified particles that only one 5.5kb  $\Delta$ Ad.AAV1 genome is packaged per capsid. This is confirmed by electron-microscopy of  $\Delta$ Ad.AAV1 particles, which reveals  
15 only spotted electron-dense staining associated with viral cores and extended free luminal space (see Figure 2).

293 cells are transfected with plasmids expressing Rep 68/78 under the CMV promoter and the kinetics of rep expression is determined by Western Blot with cell lysates  
20 collected at different time points after transfection. Next, these 293 cells are infected with Ad.AAV (MOI 1, 10, 100 pfu/cell) at specific time points after transfection of Rep plasmids depending on the Rep expression kinetics (e.g. 3, 6, 12, 24 . . . hours after transfection). It is important to time Ad.AAV infection exactly because viral DNA replication must be taking place or finished before Rep production reaches peak levels.  
25 In general, adenovirus DNA replication in 293 cells (infected with MOI 10) is maximal at 18 hours post-infection, followed by production of structural proteins, packaging of viral genomes, and breakdown of cellular membrane structures (which is concluded ~36-48h p.i.) (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; van der Vliet, B., 1995, *In* w. Doerfler, et al. (eds.) vol. 2  
30 p. 1-31, Springer-Verlag, Berlin). Viruses are collected 48h after infection and banded by CsCl ultracentrifugation. Viral material from purified bands corresponding to



ΔAd.AAV is lysed, DNase-treated (to liberate DNA associate Rep) and subjected to immunoprecipitation-Western Blot with Rep specific antibodies to detect co-packaged Rep. Based on theoretical calculations assuming that two Rep molecules bind per Ad genome, ~1-10ng Rep proteins is expected from Lysates of  $10^{10}$  particles, which is within the range of detectability by Western Blot. Alternatively, co-packaged Rep may be detected based on its functional activity to mediate rescue and site-specific integration of the AAVITR transgene cassette. To test whether functional Rep protein is co-packaged into hybrid vector particles, CsCl purified ΔAd.AAV1 particles generated in 293 cells co-expressing Rep after Ad/AAV1 infection (ΔAd.AAV1+Rep) can be used for transduction studies. Three days after ΔAd.AAV1+Rep infection of the human cell line K562, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE. If efficient Rep-mediated site-specific integration of excised AAV cassettes is successful, then other ΔAd.AAV+Rep hybrid vectors with β-Gal and SNori as transgenes can be produced.

Integration studies with Rep vectors in erythroid cells:

The hypotheses behind the rational of a rep-expressing hybrid vector (ΔAd.AAV-rep) are: (1) transient Rep co-expression from ΔAd.AAV-rep vectors can enhance site-specific vector integration in human cells and (2) integration occurs via the AAV ITR(s) without the rep gene, which is placed outside the AAV cassette, thus eliminating rep expression upon vector integration. To test hypothesis 1, transduction frequencies of ΔAd.AAV/rep versus ΔAd.AAV vectors can be compared based on the formation of G418 resistant colonies and quantify site-specific integration events at different time points after infection of human and mouse cells by PFGE and PCR. To test hypothesis 2, the structure of integrated vector in transduced cell populations and single clones can be delineated by Southern analysis and by sequencing of vector/chromosomal DNA junctions. These studies can be performed with ΔAd.AAV-rep, ΔAd.AAV, and ΔAd.AAV+Rep (copackaged protein) in human K562 or HEL (for AAVS1 integration) and mouse MEL cell lines.

Cells infected with  $\Delta$ Ad.AAV-SNori,  $\Delta$ Ad.AAV-SNori+Rep or  $\Delta$ Ad.AAV-Snori-rep can be subjected to G418 selection. The number of G418 resistant colonies determined after 4 weeks of selection in relation to the number of initially infected cells. The selection process for colonies that did not survive continued selection due to potential rep-mediated cytotoxicity or episomal vector expression can be monitored. If rep expression from  $\Delta$ Ad.AAV-SNori rep does not affect cell viability and proliferation, then more G418 resistant colonies should appear in  $\Delta$ Ad.AAV-SNori-rep and  $\Delta$ Ad.AAV-Snori+Rep. The structure of integrated vector can be determined by Southern Blot and sequencing of integration junctions.

To uncover a potential selection bias against rep producing cells after transduction with  $\Delta$ Ad.AAV/rep, site-specific and random vector integration events can be quantitated in cellular DNA isolated from cell populations at different time points after infection (e.g. 0.5, 1, 3, 7, 14 days). To do this, the techniques based on PFGE-Southern can be utilized.

It is expected that the signal(s) for AAVS1-specific integration in  $\Delta$ Ad.AAV/rep infected human cells increases during the first days after infection and then remains constant over time.

In a separate study, the integration status of vector DNA (analyzed by PFGE or PCR) and the number of integrated copies (analyzed by Southern Blot) with the expression level of  $\beta$ -galactosidase in single clones transduced with  $\beta$ -Gal hybrid vectors ( $\Delta$ Ad.AAV-BG,  $\Delta$ Ad.AAV-BG+Rep, or  $\Delta$ Ad.AAV-BG-rep) can be correlated. Together with data obtained in the studies described in the Specification, this allows assessment of whether transcriptional silencing is associated with site-specific vector integration into the AAVS1 site.

It is not clear a priori whether the specific Rep function for vector rescue, concatemerization, and integration can efficiently occur in non-S-phase or non-dividing cells. To test whether  $\Delta$ Ad.AAV<sup>fx</sup>,  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV-rep/+Rep vectors can integrate into non-dividing cells, transduction studies in cell cycle arrested cell cultures can be performed as described earlier.

## DISCUSSION

5 The establishment of stable cell lines expressing Rep 68/78 at detectable levels is not possible, which is probably due to rep mediated cytotoxicity. Therefore, it is not possible to perform long-term transduction studies (e.g. G418 selection or studies in single clones) in combination with ectopic rep expression. Moreover, due to the inhibitory effect of rep on adenovirus replication, it is currently not possible to generate adenoviral vectors expressing rep under the RSV or PGK promoter.

10

Taken together, this indicates that co-expressed Rep may stimulate site-specific transgene integration.

### 15 E. A Detailed Study Of Transduction/Integration Of Hybrid Vectors In Erythroid Cell Lines:

In order to improve transduction and integration frequencies of the hybrid vectors into erythroid cell lines, a detailed study comparing various hybrid vectors have to be carried out as described below. The transduction studies are performed in K562 cells which is  
20 considered to be an adequate model to study gene transfer vehicles into erythroid cells (Floch, V., et al., 1997, *Blood Cells, Mol. and Diseases*.23, 69-87). The optimal vectors should be able to integrate into the cellular genome with a high frequency, determined by Pulse field gel electrophoresis (PFGE) and Southern blot as described in Example 4. In addition, the results from the following studies will serve to evaluate whether a given  
25 hybrid vector needs to be modified for site-specific integration in the host genome.

#### Sequencing of integration junctions:

30 The ultimate proof for vector integration is the sequencing of junctions between SNori vector DNA and chromosomal DNA. Furthermore, this clarifies the question whether the AAV ITRs represent the substrate for integration. Specifically, DNA from clones with

known  $\Delta$ Ad.AAVSNori integration structure (analyzed by Southern Blot) digested with EcoRI, which does not cut within the SNori cassette. The resulting fragments are circularized and transformed into a specific *E. coli* strain (according to the protocol described by Rutledge and Russell (Rutledge, E. A. et al., 1997, *Journal of Virology*, 5 71:8429-36)). Kanamycin resistant bacterial clones should contain the integrated SNori cassette. Flanking chromosomal DNA in rescued plasmids can be sequenced with primers specific to the transgene.

To confirm vector integration in a small number of transduced cells, genomic DNA is 10 extracted and digested with EcoRI. EcoRI fragments are ligated to linkers containing a specific primer binding site and are then digested with NotI, religated and propagated in *E. coli*. Plasmid DNA from a representative number of bacterial clones is sequenced to determine the vector/chromosomal DNA junctions.

15 Dose dependent toxicity:

In order to test that the transduction frequency is dose-dependent and  $\Delta$ Ad.AAV vectors, which are devoid of all adenoviral genes, could be used to infect cells at higher doses with less cytotoxicity than first generation adenovirus, K562 cells are infected with 20 different MOIs ( $1-10^8$ ) of  $\Delta$ Ad.AAVBG and the first generation vector Ad.AAVBG (which contains the same  $\beta$ -Gal expression cassette). At day 4 post-infection, the total number of cells, the percentage of viable cells (based on trypan blue exclusion) and the percentage of X-Gal positive cells are counted. A fraction of infected cells are quantified for  $\beta$ -Gal expression using the Galacto-Light kit. The level of transgene expression is 25 expected to be comparable between the two vectors. K562 cells are predicted to tolerate higher doses of  $\Delta$ Ad.AAVBG better than Ad.AAVBG which express viral genes.

Integration frequency with and without G418 selection:

30 In order to investigate the integration frequency of the different vectors and to confirm that AAV ITRs present in double-stranded adenoviral DNA genomes can mediate vector

integration with a frequency comparable to rAAV vectors. integration studies are performed based on the formation of G418 resistant colonies with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.SNoriITR, and rAAVSNori after infection with  $2 \times 10^5$  and  $2 \times 10^6$  genomes per cell (Fig. 8). After infection, cells are plated in 96 well plates under limiting dilution and selected with G418 to estimate the frequency of formation of G418 resistant colonies. Another set of cells is plated without G418. A representative number of clones (w/ and w/o G418 selection) are expanded to  $>10^6$  cells (after 3-4 weeks of culture) and analyzed for the presence of viral DNA by Southern Blot as well as PFGE analysis to discriminate between episomal vector DNA and vector genomes stably associated with chromosomal DNA. This allows us to estimate the integration frequency of the different vectors, to assess the effect of G418 selection on integration, and to consider position effects on neo expression in calculating the total integration frequency. Integrated vector copies with a frequency of at least  $1 \times 10^{-4}$  is predicted only for  $\Delta$ Ad.AAVSNori and rAAVSNori. The total number of colonies may be lower in both the first generation vectors, Ad.AAVSNoriITR and Ad.SNori, due to the toxic effects of expressed adenoviral proteins; however, a higher integration frequency is predicted for the vector containing the AAV ITRs (Ad.AAVSNoriITR).

#### Kinetics of integration:

Compared to rAAV, the double-stranded nature of entering  $\Delta$ Ad.AAV genomes provides more protection against degradation. Furthermore, the synthesis of transcriptionally active double-stranded intermediates from single-stranded genomes, which is considered a limiting step in rAAV transduction, is not required in  $\Delta$ Ad.AAV transduction. Thus, the lag phase between infection and expression seen with rAAV vectors, which is causally linked to double-strand synthesis/integration may be shorter or absent in infections with  $\Delta$ Ad.AAV vectors. Furthermore, it was demonstrated earlier that a 9kb mini-adenoviral genome packaged into adenoviral particles is only short lived and completely degraded by day 3 post-infusion. In contrast, transduction with the 5.5kb  $\Delta$ Ad.AAV1 (Figure 1) genome allows for long-term expression, suggesting that either

AAV ITRs can stabilize the viral genome as an episome until it is integrated or integration occurs shortly after infection.

The status of vector DNA can be examined in K562 cells at different time points after infection with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.AAVSNoriITR, or rAAVSNori (MOI 5  $2 \times 10^5$ ). Infected cells are harvested at 1 hour, 5 hours, 1 day, 3, 7, and 14 days after infection and chromosomal DNA is analyzed by PFGE followed by hybridization with a transgene specific probe. This technique allows us to distinguish between episomal vector DNA, which appears as a distinct 5.0kb band and integrated DNA. Furthermore, 10 extra chromosomal high-molecular weight vector concatemers can be detected. In the case of random integration, after digestion of chromosomal DNA with I-CeuI or PI-SceI, vector-specific signals in the range of 1-2mb should be seen. The intensity of episomal and integrated vector signal is quantified for each time point using phosphorimager analysis. This gives information about the kinetics of hybrid vector integration in a 15 population of infected K562 cells and the intracellular stability of hybrid vector genomes.

Structure of integrated vector DNA and integration junctions with chromosomal DNA:

$\Delta$ Ad.AAV1 integrates as concatemer/s randomly into host DNA as shown previously. 20 How many vector copies are present in one concatemer and whether the extent and the kinetics of tandem-formation are dose dependent still remain unclear. Another unanswered question is how  $\Delta$ Ad.AAV integrates: whether one or both ITRs are involved, whether the integrated ITRs are still intact, and whether adenoviral sequences integrate as well. These issues are important for the strategy to include rep genes into the 25 hybrid vector genome. Moreover, if intact AAV ITRs are present within integrated vector copies, helper virus (adenovirus or HSV) infection in vivo may mobilize the integrated AAV-ITR vector cassette and affect stability of transgene expression.

To answer these questions, K562 cells can be infected with  $\Delta$ Ad.AAVSNori, AdSNori, 30 Ad.AAVSNoriITR, and rAAVSNori at MOIs  $2 \times 10^5$ ,  $2 \times 10^6$ , or  $2 \times 10^7$  genomes per cell. Infected cells are plated in 96 well plates in the presence or the absence of G418. The

latter is included because G418 may cause amplification of integrated vector DNA (Rutledge, E. A. et al., 1997, *Journal of Virology*, 71:8429-36). Genomic DNA from isolated clones can be analyzed by regular Southern Blot as described in the Examples Section to confirm the presence of vector concatemers and calculate the number of integrated vector copies. More informative is the sequencing of integrated vector copies and their junctions with chromosomal DNA. The structure of integration junctions can be delineated using the role of AAV ITRs in vector integration and the extent of insertional mutagenesis after transduction. This data provides information about the potential risks of hybrid vector used in clinical trials.

Transduction of cell cycle arrested cells:

The ultimate target for the hybrid vectors described in the Specification are quiescent hematopoietic stem cells. We hypothesize that the double-stranded nature of  $\Delta$ Ad.AAV genomes and specific nuclear import mechanisms may allow for the transduction of non-dividing cells. This is in part supported by the transduction studies with  $\Delta$ Ad.AAV1 in quiescent hepatocytes in vivo. To confirm this data, primary fibroblasts can be forced to enter the G<sub>0</sub> phase by serum/growth factor starvation before infection with the hybrid vectors according to a protocol described by Russell (Russell, D. et al., 1995, *PNAS*, 92:5719-23). Cells are maintained for three days after infection under serum/growth factor deprivation. At this time point, genomic DNA is isolated and analyzed for integration events by PFGE in comparison with growing cells. Another series of integration studies can be carried out on K562 cells arrested in the G<sub>1</sub>/S phase of the cell cycle with aphidicolin (added 1 day before and maintained several days after infection with hybrid vectors - depending on the integration kinetics studies described earlier). To investigate whether DNA damaging agents increase the transduction frequency of hybrid vectors, cell-cycle-arrested K562 cells or primary fibroblasts can be treated with cisplatin or <sup>3</sup>H-thymidine prior to virus infection according to a protocol described by Alexander and Russell (Alexander, I. E. et al., 1994, *J. Virol.*, 68:8282-87; Russell, D. et al., 1995, *PNAS*, 92:5719-23). Furthermore, the effect of chromosomal DNA decondensation on the transduction efficiency of hybrid vectors can be studied in arrested

cells after treatment with puromycin, staurosporin, Hoechst 3328, distramycin, or vandate.

#### F. Improvements in $\Delta$ Ad.AAV Production and Purification

5 To inhibit packaging of full-length genomes a modified form of I-Sce I, a yeast mitochondrial intron-endonuclease with a non-palindromic 18-bp recognition sequence is expressed in 293 cells. Constitutive expression of this enzyme in mammalian cells is not toxic, possibly due to either the lack of I-SceI sites in the genome or sufficient repair of  
10 them (Rouet P. et al, 1994, *PNAS*, 91:6064-8). The yeast I-Sce I is modified with an SV40 T-antigen nuclear localization signal and an optimal Kozak sequence to enhance its functionality in mammalian cells (Rouet P. et al, 1994, *PNAS*, 91:6064-8). For another yeast endonuclease it was shown that a recognition site within an transduced Ad genome was efficiently (30% of all transduced genomes) when expressed in human A549 cells.  
15 Importantly, the expression of E4 ORF6 and ORF3 expressed from the transduced Ad genome inhibited double-strand break repair mediated by the endonuclease (Nicolas, A. L. et al, 2000, *Virology*, 266:211-24). This is consistent with the observations by others where these E4 proteins prevent concatemerization of the viral genome (Boyer, J. et al, 1999, *Virology*, 263:307-12). Based on this, packaging of full-length virus containing a  
20 I-SceI recognition site is reduced in 293 cells constitutively expressing I-Sce I. The 18mer I-Sce site is inserted into the E3 region of the Ad.IR vectors. These vectors are generated and amplified in 293 cells followed by a large-scale infection of 293 cells expressing I-SceI. Alternatively, an expression cassette for the endonuclease XhoI is inserted into the E3 region of Ad.IR or Ad.AAV vectors. The XhoI gene will be  
25 modified for optimal function in mammalian cells. Vectors expressing XhoI are generated and amplified in 293 cells expressing the Xho I isoschizomer PaeR 7 methyltransferase (PMT) (Nelson, J. E. et al, 1997, *J. Virol.*, 71:8902-7), which mediates the addition of a methyl group onto the N6 position of the adenine base of Xho I sites, CTCGAG. This protects the viral and cellular genome from XhoI cleavage. Methylated  
30 Ad vectors are produced at high titers.  $\Delta$ Ad.AAV vectors are then obtained by large-scale infection of 293 cells with the Ad.AAV-XhoI vectors. At this stage the viral



genome is not methylated and is digested at the XhoI sites. XhoI sites present within the transgene cassette are deleted by site-directed mutagenesis without altering the amino acids sequence. (XhoI is accumulated only at late stages in virus replication and should act only upon a large part of Ad DNA when replication is completed. In addition, ultracentrifugation optimizes the separation between  $\Delta$ Ad.IR and  $\Delta$ Ad.IR particles (Blague, C. et al., 2000, *Blood*, 95:820-8).

## EXAMPLE II

### 10 MODIFIED FIBER PROTEIN

#### A. Test the Infectivity of Different Human or Animal Serotype on Human Bone Marrow Cells.

15 Since the amino acid sequence of the fiber knob region varies considerably among the ~50 known serotypes, it is thought that different adenovirus serotypes bind to different cellular receptor proteins or use different entry mechanisms (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14; Defer, M., et al., 1993, *J. of*  
20 *Virology*, 64, 3661-3673). Although most adenoviruses contain RGD motifs in the penton base proteins, there are a number of serotypes (e.g. Ad 40, 41) without this conserved sequence. These types may use integrin  $\alpha$ v-independent pathways for virus internalization (Davison, A.J., et al., 1993, *J. Mol. Biol.*, 234, 1308-1316; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14). To test whether other Ad serotypes can  
25 infect stem cell subpopulation present in human bone marrow, studies with a series of different human Ad serotypes and animal viruses can be performed (see Table II). As a means to verify efficient transduction with Ad serotypes, viral DNA is tagged before infection and the presence of viral genomes in the nuclei of transduced cells is investigated. Furthermore, whether viral DNA is replicated in transduced cells can be  
30 analyzed as indirect proof for early viral gene expression. A direct detection of expressed viral proteins is impossible due to the unavailability of antibodies against all the serotypes included in this study. Simultaneously with the infection assay, transduced

human bone marrow cells can be analyzed for morphological and immunohistochemical features characteristic of HSC or progenitor subpopulations. For retargeting, serotypes which are able to infect CD34+ subsets of bone marrow cells at the lowest MOI are selected. As the next step, the fiber gene is PCR-cloned from serotypes with potential  
5 HSC/CD34+ tropism and inserted into standard shuttle plasmids for Ad5 vector generation replacing the Ad5 fiber gene using an E. coli recombination system (Figure 9).

## METHODS

10

### Cells and viruses:

HeLa (human cervix carcinoma, ATCC CCL-2.2), CHO (chinese hamster ovary, ATCC CCL-61), K562 (human hematopoietic, ATCC 45506), HEp-2 (human larynx carcinoma,  
15 ATCC CCL-23), 293 (human embryonic kidney, Microbix, Toronto Canada) cells were maintained in DMEM, 10% FCS, 2 mM glutamine, and Pen/Strep. Culture media for CHO cells was supplemented with 200μM asparagine and 200μM proline. Human CD34+-enriched bone marrow cells were purified from peripheral blood after mobilization using MiniMACS VS<sup>+</sup> separation columns (Miltenyi Biotec, Auburn, CA),  
20 according to the manufacturer's instructions. Aliquots were stored in liquid nitrogen. Sixteen hours before the experiment, cells were recovered from the frozen stock and incubated overnight in IMDM media, supplemented with 20% FCS, 10<sup>-4</sup> M β-mercaptoethanol, 100 μg/ml DNaseI, 2 mM glutamine, 10 U/ml IL-3, and 50 ng/ml stem cell factor (SCF) or 2 ng/ml thrombopoietin (Tpo). The purity of CD34+ preparations  
25 was verified by flow cytometry and was consistently greater than 90%.

### Flow cytometry:

Adherent cells (CHO, HeLa) grown in non-tissue culture treated 10 cm dishes (Falcon,  
30 Franklin Lakes, NJ) were detached by treatment with 1mM EDTA and washed three times with wash buffer (WB), consisting of PBS supplemented with 1% FCS. Cells

grown in suspension (K562, CD34+) were washed three times with WB. After washing, cells were resuspended in WB at  $2 \times 10^6$  cells/ml.  $2 \times 10^5$  cells were incubated in WB for 1 h at 37°C with monoclonal antibodies specific for  $\alpha_v$ -integrins [L230, ATCC: HB-8448, (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795) (1/30 final dilution),  
5 CAR [RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652) (1/400 final dilution)], or BrdU [(Amersham, Arlington Heights, IL) (1/100 final dilution)]. Subsequently, cells were washed with WB, and incubated with fluorescein isothiocyanate (FITC)-labeled horse anti-mouse IgG antibodies [(Vector Labs., Burlingame, CA) (1/100 final dilution)] or phycoerythrin (PE)-  
10 labeled goat anti-mouse IgG antibodies [(Calbiochem, La Jolla, CA) 1:100 dilution] for 30 min at 4°C. After incubation with secondary antibodies, cells were washed two times with WB and  $10^4$  cells per sample were analyzed in duplicate by flow cytometry.

For the analysis of CD34 and c-kit expression on transduced CD34+ cells and for  
15 fluorescent activated cell sorting (FACS), purified human CD34+ cells were incubated with phycoerythrin(PE)-conjugated anti-CD34 monoclonal antibodies (Becton-Dickinson Immunocytometry Systems, San Jose, CA) or with PE-labeled anti-CD117 (c-kit) monoclonal antibodies (MAb 95C3, Immunotech, Beckman Coulter, Marseille, France) according to the manufacturer's protocol followed by flow cytometry analysis. All  
20 analyses and sortings were performed on a FACStar Plus flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with 488nm argon and 633 nm HeNe lasers. For analysis of c-kit expression and FACS purification of CD34+/c-kit+ cells, SCF was not added to the media during culturing of CD34+ cells.

## 25 RESULTS

### CAR/ $\alpha_v$ -integrin expression on test cells:

It is generally accepted that CD34+ cells possess bone marrow repopulating activity.  
30 Therefore, we used human CD34+ cells as the target for our studies towards identifying Ad serotypes with HSC tropism and constructing new viral vectors. Studies were

performed on mobilized, CD34-positive, peripheral blood cells from one donor under conditions which are known to retain CD34+ cells in a quiescent stage (Leitner, A., et al. 1996. *Br.J.Haematol.* 92:255-262; Roberts, A. W., Metcalf, D. 1995. *Blood.* 86:1600-1605). More than 90% of purified cells were CD34 positive by flow cytometry.

5 Furthermore, we included into our Ad tropism studies the cell line K562, which is considered to be an adequate model for studying gene transfer into human hematopoietic cells (McGuckin, et al. 1996. *British Journal of Haematology.* 95:457-460). HeLa cells, which are readily infectible by Ad5, and CHO cells, which are refractory to Ad5 infection (Antoniou, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9), were used as positive and

10 negative control cell lines, respectively.

For Ad5, both, binding to the primary receptor and to  $\alpha_3\beta_5$  and  $\alpha_v\beta_5$  integrins are important for high efficiency infection of target cells. The expression of CAR and  $\alpha_v$  integrins on test cells was analyzed by flow cytometry using monoclonal antibodies

15 against CAR (RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652)) and  $\alpha_v$  integrins (L230 (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621)) (Figure 10). As expected, nearly all HeLa cells expressed high levels of CAR and  $\alpha_v$ -integrins, whereas CHO cells lacked significant CAR and  $\alpha_v$ -integrin expression. Fifteen and 77% of K562 cells expressed CAR and  $\alpha_v$ -

20 integrins, respectively. Only ~6% of the CD34+ cells used in our studies expressed CAR and 17% were positive for  $\alpha_v$ -integrins. Notably, the preparation of CD34+ cells represents a mixture of different cell types. The absent or low expression of primary and secondary Ad5 receptors on non-cycling human CD34+ cells is in agreement with previous reports (Huang, S., et al. 1996. *J. Virology.* 70:4502-4508; Neering, S. J., et al.

25 1996.. *Blood.* 88:1147-1155; Tomko, R. P., et al.. 1997. *Proc. Natl. Acad. Sci. USA.* 94:3352-3356).

Infection assay using wild-type Ad5 and K562 cells:

30 The presence of viral DNA in the nucleus of infected cells is an indirect means to demonstrate efficient virus binding, internalization, and nuclear import. Nuclear

localization of the viral genome is a prerequisite for transgene transcription and integration. Two techniques are utilized to tag viral DNA for in situ analysis. To optimize the infection assay, wild-type Ad5 virus and K562 cells which are permissive for Ad5 infection can be used. The first protocol (Challberg, S.S. and Ketner, S. 1981, *Virology* 114, 196-209), is based on  $^{32}\text{P}$ -labeling of viral DNA. During amplification of wild-type Ad5 and A549 cells,  $^{32}\text{P}$ -phosphate (40 $\mu\text{Ci/ml}$ ) is added to phosphate-free medium. After development of CPE,  $^{32}\text{P}$ -tagged virus is harvested, banded in CsCl gradients, and titered on HeLa cells according to standard protocols. To simulate the conditions for infection of human bone marrow cells, K562 cells are incubated in suspension with a MOI of 1, 10, or 100 of  $^{32}\text{P}$ -Ad5 for 2, 4, 6, or 8 hours under agitation at 37°C. This covers the time period necessary for adsorption, internalization, and nuclear import. After washing, cells are fixed either transferred to microscopy slides using cytopsin or embedded in paraffin and sectioned (according to protocols from VECTOR labs, Burlingame, CA). The latter has the potential advantage that multiple consecutive sections (5 $\mu\text{m}$ ) of the same cell can be analyzed by different methods (e.g. for  $^{32}\text{P}$  tagged viral DNA, for specific histological staining, for immunofluorescence), which allows for correlating infection with a particular cell type present in the bone marrow. Cells are incubated in a Kodak NTB-2 photo emulsion for autoradiography. The exposure time can be optimized to minimize background or non-nuclear localized signals. A dose and time dependent appearance of nuclear silver grains is expected under the optimized conditions. Since  $^{32}\text{P}$ -phosphate can label viral proteins as well, a cytoplasmic background signal might appear. To facilitate detection, immunofluorescence with HSC specific antibodies on sections can be performed. As an alternative method, a BrdU-labeling technique for viral DNA can be used (Lieber, A., et al. 1999. *J Virol* 73:9314-24; Lieber, A. et al., 1996, *Journal of Virology*, 70:8944-60). In this case, different amounts of BrdU are added to the A549 culture medium during wtAd5 virus propagation. BrdU labeled viral DNA can be detected with monoclonal antibodies specific to BrdU. The signal can be enhanced using layers of species-specific polyclonal antibodies in combination with biotin/avidin and a fluorescent marker. BrdU tagged viral DNA can be detected on cytopsin of bone marrow cells together with cell surface markers by double or triple immunofluorescence.

## DISCUSSION

The interaction of selected Ad serotypes with CD34+ cells was tested. As a result of this screening we constructed a first-generation, Ad5-based vector whose fiber was substituted with the fiber derived from Ad35. We demonstrated that this capsid modification allowed for efficient viral transduction of potential HSCs by the corresponding chimeric Ad vectors.

All tropism and transduction studies were performed with non-cycling CD34+ cells, which are thought to include HSCs. The quiescent stage of CD34+ cells purified from mobilized blood is important because induction of cell proliferation is associated with a loss of the ability to reconstitute hematopoiesis and with changes in the spectrum of cellular receptors (Becker, P. S., et al. 1999. *Exp. Hematol.* 27:533-541). It is known that treatment of hematopoietic cells with cytokines or growth factors changes the expression of specific integrins including  $\alpha_v$ -integrins, which would ultimately alter the susceptibility of cells to Ad infection or may effect viability of infected cells (Gonzalez, R., et al. 1999. *Gene Therapy*. 6:314-320; Huang, S., et al. 1995. *J. Virology*. 69:2257-2263). Another fact that complicates the interpretation of transduction studies is the extraordinary heterogeneity of CD34+ cells in regards to morphology and function.

### B. Screening different adenoviruses to establish tropism to HSC.

The ATCC provides more than 70 different human or animal adenoviruses (see Appendix I). A collection of 15 human serotypes and 6 animal adenoviruses (see Table II) are selected based on the following criteria: (i) availability of the complete genome sequence or fiber sequence from the NIH gene bank (ii) CAR receptor usage absent or unknown, (iii) different subgroups, and (iv) moderate or low tumorigenicity (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia).

However, any serotype shown in the Appendix hereto can be used for the invention described. Animal viruses are included in the infectivity assay because this may provide a

means to circumvent the pre-existing humoral immunity against human Ad5 fiber, which represents a critical obstacle for clinical trials with Ad vectors.

## METHODS

5

### Viruses:

The following human adenovirus serotypes were purchased from the ATCC: 3 (VR-3), 4 (VR1081), 5 (VR-5), 9 (VR1086), 35 (VR-716) and 41(VR-930). Adenovirus No. VR-  
10 716 was purchased from ATCC labeled as serotype 34, however it was found to be serotype 35 upon sequencing of the fiber region. For amplification, the corresponding Ads were infected onto HeLa, 293, or HEp-2 cells under conditions that prevented cross-contamination. Virus was banded in CsCl gradients, dialyzed and stored in aliquots as described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).  
15 Plaque titering was performed as follows: Confluent 293 cells plated in 6-well plates were incubated for 24 hours with virus in a total volume of 1ml. Two weeks after infection, plaques were counted on cultures overlayed with 1% agarose/MEM/10% FCS.0

### 20 EM studies:

CsCl-banded Ad stocks were thawed and diluted with 0.5% glutaraldehyde. Grids were prepared as described earlier (Mittereder, N., et al. 1996. *J. Virology*. 70:7498-7509). After staining with 2% methylamine tungstate (Nanoprobes, Stony Brook, NY), the  
25 carbon-coated grids were evaluated and photomicrographed with a Phillips 410 electron microscope, operated at 80 kV (final magnification 85,000x). For each particular Ad serotype, the number of morphologically deficient viral particles per 100 was counted in five random fields.

## 30 RESULTS

Electron microscopy:

Little is known about the stability of particles from serotypes other than Ad5. Since the intactness of viral particles was crucial for comparative interaction studies, virions from the serotypes specified above were analyzed by electron microscopy (EM). EM studies of negative contrast stained Ad suspensions demonstrated that the percentage of defective particles (loss of icosahedral shape or luminal staining) did not exceed 5% indicating that serotype preparations had comparable qualities. Representative EM photographs are shown for Ads 5, 9, and 35 (Figure 11).

Serotype screening:

It is thought that different Ad serotypes bind to different cellular receptor proteins and use different entry mechanisms (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Mathias, P., et al. 1994. *Journal of Virology*. 68:6811-6814). A set of human adenoviruses was obtained from the ATCC to be tested for tropism to CD34+ cells. These included serotypes 3, 4, 5, 9, 35, and 41 representing different subtypes (Table 1). We believed that these serotypes would use different cellular attachment and internalization strategies due to differing lengths of fiber shafts (Chroboczek, J., et al. 1995. *Adenovirus fiber*, p. 163-200. In a. P. B. W. Doerfler (ed.), *The molecular repertoire of adenoviruses*, vol. 1. Springer Verlag, Berlin; Roelvink, P. W., et al. 1998. *J. Virology*. 72:7909-7915), the presence or absence of RGD motifs within the penton base, and differing tissue tropism. The relatively little characterized Ad35 was selected because it was found in immunocompromised hosts, particularly in bone marrow recipients (Flomenberg, P., et al. 1994. *Journal of Infectious Diseases*. 169:775-781; Flomenberg, P. R., et al. 1987. *Journal of Infectious Diseases*. 155:1127-1134; Shields, A. F., et al. 1985 *New England Journal of Medicine*. 312:529-533). The latter observations prompted us to believe that bone marrow cells are among the natural reservoirs for Ad35.



TABLE II

Human and animal adenoviruses with potential interest for the invention

| Adenovirus | Human/<br>Group B                   | Human/<br>Group<br>D         | Human/<br>Group<br>F  | Avian        | Bovine | Canine | Ovine | Swine | Mouse |
|------------|-------------------------------------|------------------------------|-----------------------|--------------|--------|--------|-------|-------|-------|
| Serotype   | <u>3</u> , 7, 11, 16,<br>21, 34, 35 | 8, 15, 17,<br>19, 28, 3<br>7 | <u>40</u> , <u>41</u> | CELO,<br>EDS | 3      | 1, 2   | 5     | 4     | 1     |

The underlined serotypes use CAR independent pathways for cell entry.

- 5 For amplification, the corresponding adenovirus stocks can be infected onto HeLa or A549 cells such that at a given time only one virus type is handled in a separate laminar flow hood and cultured in Hepa-filtered bottles, preferentially in separate CO<sub>2</sub> incubators to avoid cross-contamination. During propagation, viral DNA is tagged using one of the techniques described earlier. Viral DNA can be isolated from purified particles. The
- 10 XhoI restriction pattern is analyzed for methylated and unmethylated viral DNA by Southern blot using the full genome of the corresponding virus type as a radioactive probe.

#### DISCUSSION

15

- Although it was reported earlier by slot-blot assay that fiber knobs derived from 2, 9, 4, and 41L can bind to CAR (Roelvink, P. W., et al., 1998. *J. Virology*. 72:7909-7915), it is not clear whether this binding occurs with an affinity that is physiologically relevant and whether this would confer cell entry. Furthermore, as shown for the Ad5 interaction
- 20 between the penton and integrins, a secondary receptor is required to induce virus internalization. We demonstrated that different serotypes interacted differently with the K562 or CD34+ target cells. Ad5, Ad4, and Ad41 were not able to efficiently attach to and be internalized by K562 and CD34+ cells. Although Ad4 belongs to a separate subgroup (E), it is thought that Ad4 represents a natural hybrid between subgroup B and
- 25 C viruses with a fiber related to Ad5 (Gruber, W. C., et al. 1993. *Virology*. 196:603-611).

Therefore, it was not surprising that Ad4 has binding properties similar to Ad5. The subgroup F serotype Ad41 has been shown to contain distinct fibers, a long shafted and a short-shafted fiber allowing for different cell entry pathways (Tiemessen, C. T., Kidd, A.H. 1995. *J. Gen. Virol.* 76:481-497). The Ad41 penton base does not contain RGD motifs suggesting that this virus may use  $\alpha_v$ -integrin independent pathways for cell entry. However, these features did not improve interaction with CD34+ cells. Ad9, Ad3, and Ad35 did interact with CD34+ cells more efficiently than Ad5. Out of all the serotypes tested, Ad35 demonstrated the most efficient attachment and internalization with K562 and CD34+ cells. Although the short-shafted Ad9 can bind to CAR, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Therefore, the low level of  $\alpha_v$ -integrin expression on certain subset/s of CD34+ cells may account for the observed susceptibility to Ad9.

#### C. Attachment and Internalization of the Ad serotypes to K562 and CD34+ cells.

##### *METHODS*

##### Labeling of Ads with [ $^3$ H]-methyl thymidine:

Serotypes were labeled with [ $^3$ H]-methyl thymidine as described in detail elsewhere (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Briefly,  $5 \times 10^7$  HeLa or 293 cells were grown in 175 sq. cm flasks with 15 ml DMEM/10% FCS and infected with wild type adenovirus at a MOI of 50 or higher. Twelve hours post-infection, 1 mCi of [ $^3$ H]-methyl thymidine (Amersham, Arlington Heights, IL) was added to the media and cells were further incubated at 37°C until complete CPE was observed. Then, cells were harvested, pelleted, washed once with cold PBS, and resuspended in 5 ml PBS. Virus was released from the cells by four freeze-thaw cycles. Cell debris was removed by centrifugation and viral material was subjected to ultracentrifugation in CsCl gradients and subsequent dialysis as previously described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology.* 70:8944-8960). Virus purification and dialysis removed unincorporated radioactivity. Wild type Ad particle concentrations were determined

spectrophotometrically by measuring the OD<sub>260</sub>, utilizing the extinction coefficient for wild-type Ad5  $\epsilon_{260} = 9.09 \times 10^{13}$  OD ml cm virion<sup>-1</sup> (Maizel, J. V., et al. 1968. *Virology*. 36:115-125). The virion specific radioactivity was measured by a liquid scintillation counter and was always in the range of  $1 \times 10^5$  to  $1 \times 10^4$  cpm per virion. For selected  
5 variants, the fiber gene was PCR amplified and sequenced to ensure identity and the absence of cross-contamination.

Viral DNA tagged with methylase and test for replication by genomic Southern blots:

10 To ultimately confirm transduction, a protocol to detect adenoviral replication in infected cells can be established. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. A site-specific methylation strategy is utilized to monitor viral DNA replication within infected cells (Nelson, J. et al., 1997, *Journal of Virology*. 71:8902-07). Methylation marked adenovirus can be produced by the addition of a  
15 methyl group onto the N6 position of the adenine base of XhoI sites, CTCGAG, by propagation of the virus in HeLa or A549 cells expressing the XhoI isoschizomer PaeR7 methyltransferase (PMT) (Kwoh, T.J., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 7713-7717). It is known that methylation does not affect vector production but does prevent cleavage by XhoI. Loss of methylation through viral replication restores XhoI cleavage  
20 and can be detected by Southern blots of genomic DNA from infected cells in comparison to native, non-methylated, viral genomes.

Attachment and internalization assays:

25 These studies were performed based on a protocol published elsewhere (Wickham, T. J., et al. 1993. *Cell*. 73:309-319). In preliminary experiments, we found that labeled Ad5 virions reached equilibrium in attachment to HeLa cells after 45 min at 4°C with an MOI of 400 pfu per cell. For attachment studies,  $3.5 \times 10^5$  cells were incubated for one hour on ice with equal amounts of [<sup>3</sup>H]-labeled adenovirus OD particles equivalent to an MOI of  
30 400 pfu/cell for Ad5 in 100 µl of ice-cold adhesion buffer (Dulbecco's modified Eagle's medium supplemented with 2 mM MgCl<sub>2</sub>, 1% BSA, and 20 mM HEPES). Next, the cells

were pelleted by centrifugation for 4 min at 1000 x g and washed two times with 0.5 ml ice-cold PBS. After the last wash, the cells were pelleted at 1500 x g, the supernatant was removed, and the cell-associated radioactivity was determined by a scintillation counter. The number of viral particles bound per cell was calculated using the virion specific radioactivity and the number of cells. To determine the fraction of internalized [<sup>3</sup>H]-labeled adenoviral particles, cells were incubated on ice for one hour with the corresponding virus, washed with PBS as described above, resuspended in 100 µl adhesion buffer, and then incubated at 37°C for 30 min. Following this incubation, cells were diluted 3-fold with cold 0.05% trypsin-0.5mM EDTA solution and incubated at 37°C for an additional 5-10 min. This treatment removed 99% of attached radioactivity. Finally, the cells were pelleted at 1500 x g for 5 min, the supernatant was removed, and the protease-resistant counts per minute were measured. This protocol minimizes the possibility that the internalization data were affected by receptor recycling (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795). Nonspecific binding of Ad particles to cells on ice was determined in the presence of 100-fold excess of unlabeled virus. This value routinely represented less than 0.1 % of viral load.

## RESULTS

### Attachment of Ad particles to target cells and internalization:

The selected serotypes were metabolically labeled with [<sup>3</sup>H]-thymidine, which is incorporated into viral DNA during replication. Adsorption and internalization can be experimentally dissociated by taking advantage of the observation that at low temperature (0-4°C) only virus cell attachment occurs, whereas internalization requires incubation at higher temperatures. The number of particles adsorbed or internalized per cell was calculated using the virion-specific radioactivity and used to quantify interaction of Ads 3, 4, 5, 9, 35, and 41 with CD34+, K562, HeLa and CHO cells (Figure 12). The serotypes varied significantly in their ability to attach to and to be internalized by the different cell lines. For Ad5, the degree of attachment to the cell lines tested correlated with the level of CAR expression. In CHO cells, which were previously shown to be refractory to Ad5

infection, the level of attachment and internalization was about 50-70 viral particles per cell. This number was hereafter assumed negative in terms of susceptibility of a given cell type for Ad5. Interaction of the other serotypes with CHO cells was not significantly higher indicating that corresponding receptor/s were absent on CHO cells. All serotypes tested interacted with HeLa cells; with Ad3 and Ad35 being the most efficient variants. The presence of distinct Ad3 and Ad5 receptors on HeLa cells was demonstrated previously (Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857). Ads 4, 5, and 41 did not bind to K562 cells. In contrast, Ad9 as well as the members of subgroup B, Ad3 and Ad35, efficiently interacted with K562 cells with Ad35 having the highest number of adsorbed and internalized particles. Compared to Ad5, about 25 times more Ad35 particles were attached and three-fourth of these were internalized by K562 cells. Viral interactions with CD34+ cells were generally weaker. Among the serotypes tested, only Ad9 and Ad35 were significantly internalized by non-cycling CD34+ cells. Internalization of Ad9 and Ad35 was, respectively, four and eight times more efficient than for Ad5 particles. The number of Ad35 virions internalized by CD34+ cells was almost half of that seen for Ad5 in HeLa cells, which can be readily infected with Ad5 based vectors.

Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into HeLa, 293, and CHO cells:

HeLa and 293 cells expressing high level of primary and secondary receptors for human adenoviruses are used as a positive control for virus attachment and internalization. As a negative control CHO cells are used. CHO cells do not express the primary adenoviral receptor at a detectable level, and are therefore refractory for adenoviral infection. For attachment studies, these adherent cell lines are detached from 10 cm dishes with PBS-EDTA solution (without Ca<sup>2+</sup> and Mg<sup>2+</sup>), washed three times with ice-cold PBS, resuspended in adhesion buffer, and incubated with viruses as described above in the Examples section. As expected, all adenoviral serotypes tested are efficiently attached to and internalized into HeLa cells (Table III) (Figure 13). Adenoviruses serotypes 3, 5, 35, 41, but not 9, are efficiently attached to and internalized by 293 cells. In contrast, poor

attachment and internalization of most adenovirus serotypes are observed with CHO cells. The level of attachment on CHO is about 50-70 virus particles per cell for adenoviruses serotypes 5 and 41, 115 virus particles per cell for adenovirus type 3 and about 180 particles per cell for adenovirus serotypes 9 and 35. For further analysis, numbers >300 viral particles per cell are assumed as positive and <70 viral particles per cell as negative in terms of susceptibility of a particular cell line for efficient adenoviral transduction.

TABLE III

10 **Comparative analysis of attachment and internalization of Ad5 and Ad9 to cell lines, expressing different amounts of CAR and  $\alpha\beta$ integrins.**

| Cell line    | CAR expression | $\alpha\beta$ -integrin expression | Ad9 (attached/ internalized) | Ad5 (attached/ internalized) |
|--------------|----------------|------------------------------------|------------------------------|------------------------------|
| HeLa         | ++             | ++                                 | 426/370                      | 550/500                      |
| CHO          | -              | ++                                 | 300/300                      | 70/50                        |
| 293          | ++             | ++                                 | 20/20                        | 1950/1750                    |
| Y79          | +++            | -                                  | 190/140                      | 1200/1100                    |
| K562         | -              | +                                  | 320/230                      | 60/50                        |
| Erythrocytes | ?              | ?                                  | 420/-                        | 68/-                         |

15 Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into human CD34+ bone marrow cells and K562 erythroleukemia cell line:

Previous studies showed that the human erythroleukemia cell line K562 can be transduced with Ad5-based adenoviral vectors at very high MOIs. As shown in Figure 14, only about 60 viral particles per cell of adenovirus serotype 5 are attached to and even fewer particles are internalized into these cells at a MOI of 400. In contrast to Ad5, about 320 viral particles per cell of Ad9 and about 1500 viral particles per cell of Ad35 are

attached to and about two-thirds of them are internalized into K562 cells (Figure 14B). Human unstimulated CD34+-enriched bone marrow cells obtained from frozen stocks are incubated overnight in growth medium without cytokine stimulation. The next day, the number of viable cells is calculated. For attachment studies, cells are washed three times  
5 with ice-cold PBS, resuspended in adhesion buffer and incubated with adenoviruses. Among the adenoviral serotypes tested, only adenovirus particle of Ad9 (about 150 viral particles per cell) and Ad35 (about 320 viral particles per cell) are able to attach to unstimulated CD34+ cells on the level, compared to Ad 5 (only 60 viral particles per cell). Four-fifths of these virus particles are able to be internalized by the cells.  
10 Interestingly, upon stimulation of CD34+ cells with GM-CSF and EPO/TPO for two weeks, attachment and internalization of Ad9 viral particles are significantly increased (up to 300 particles per cell). At the same time, the transient stimulation of cells with GM-CSF for two days could not increase the level of viral attachment to the cells.  
15 Based on the above finding that Ad35 serotype is able to attach and internalize into CD34+ cells most efficiently among several serotypes tested, serotype Ad35 was selected for further studies. As described in Appendix II, a chimeric vector (Ad5 GFP/F35) containing the short-shafted Ad35 fiber sequence in an Ad5 capsid was able to target a broad spectrum of CD34+ cells in a CAR/integrin independent manner.

20

#### *DISCUSSION*

In summary, from all the serotypes tested, Ad9, Ad3, and Ad35 demonstrated the most efficient attachment to and internalization with K562 and CD34+ cells. Based on  
25 adsorption/internalization data, Ad9 and Ad35 as representatives for subgroups D and B were selected for further tropism studies.

#### **D. Characterization of Ad vector replication in K562 and CD34+ cells.**

30 Comparative analysis of Ad5, and Ad9 and Ad34 to infect and to replicate in 293, K562 and CD34+ cells. The ability of the Ad9 fiber knob domain to recognize the same

primary receptor on the cell surface as Ad5 with comparable affinity was described earlier. Thus, the finding that Ad9 viral particles can only poorly attach to 293 cells is rather unexpected. In order to find out how the attachment and internalization data reflect the biological activity of adenoviruses of different serotypes, the stocks of Ad5, Ad9 and Ad35 are characterized in more detail by electron microscopy, plaque assay on 293 cells, and quantitative replication assay in K562 and CD34+ cells.

## METHODS

### Quantitative replication assay:

1x10<sup>5</sup> CD34+ or K562 cells were infected in 100µl of growth media with different MOIs of Ad5, 9, or 35 which had been amplified in 293 cells, expressing the XhoI DNA methyltransferase isoshizomer PaeR7 (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). After 2 hours of incubation at 37°C, the cells were centrifuged at 1000 x g for 5 min, the virus-containing medium was removed, the cells were resuspended in 100µl of fresh media, and then they were incubated at 37°C until harvesting. At 16 hours post-infection for K562 cells, or 36 h post-infection for CD34+ cells, 5 µg of pBS (Stratagene, La Jolla, CA) plasmid DNA was added as a carrier which could also be used as a loading control. Genomic DNA was extracted as described previously (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). One-fourth of purified cellular DNA (equivalent to 2.5 x 10<sup>4</sup> cells) was digested with HindIII, XhoI, or with HindIII and XhoI together at 37°C overnight and subsequently separated in a 1% agarose gel followed by Southern blot with chimeric Ad5/9 or Ad5/35 DNA probes. The chimeric probes, containing sequences of Ad5 and Ad9 (Ad 5/9) or Ad5 and Ad35 (Ad 5/35), were generated by a two-step PCR amplification using Pfu-Turbo DNA polymerase (Stratagene, La Jolla, CA) and viral DNA from purified particles as a template. The following primers were used for PCR (Ad5 sequences and nucleotide numbers are underlined): Ad5F1 – (nt: 32775-32805) 5'-GCC CAA GAA TAA AGA ATC GTT TGT GTT ATG-3'; Ad5R1 – (nt: 33651-33621) 5'-AGC TGG TCT AGA ATG GTG GTG GAT GGC GCC A-3'; chimeric Ad5/9F – (nt: 31150-31177, nt: 181-208) 5'-AAT



GGG TTT CAA GAG AGT CCC CCT GGA GTC CTG TCA CTC AAA CTA GCT  
 GAC CCA -3'; chimeric Ad5/9R – (nt: 32805-32775, nt:1149-1113) 5'-CAT AAC  
ACA AAC GAT TCT TTA TTC TTG GGC TTC ATT CTT GGG CGA TAT AGG  
 AAA AGG-3; chimeric Ad5/35F – (nt: 31150-31177, nt: 132-159) 5'-AAT GGG TTT  
 5 CAA GAG AGT CCC CCT GGA GTT CTT ACT TTA AAA TGT TTA ACC CCA-3',  
 chimeric Ad5/35R (nt: 32805-32775, nt: 991-958) 5'-CAT AAC ACA AAC GAT TCT  
TTA TTC TTG GGC ATT TTA GTT GTC GTC TTC TGT AAT GTA AG-3'.  
 Nucleotide numbers are given according to the sequences obtained from the NCBI  
 GenBank (accession No. M73260 / M29978 for Ad5, X74659 for Ad9, and U10272 for  
 10 Ad35). After the first amplification, the 968 bp-long Ad9, a 859 bp-long Ad35 DNA  
 fragments corresponding to the fiber genes, and a 876 bp-long Ad5 fragment  
 corresponding to the Ad5 E4 region (located immediately downstream of Ad5 fiber gene)  
 were purified by agarose gel electrophoresis. To generate chimeric DNA probes,  
 amplified Ad5 DNA was mixed with Ad9 or Ad35 fragments obtained during the first  
 15 step of PCR, and subjected to a second PCR amplification using Ad5/9F or Ad5/35F  
 primers and the Ad5R1 primer. The resulting Ad5/9 or Ad5/35 chimeric DNA fragments  
 (see Fig. 15) were purified and their concentrations were measured  
 spectrophotometrically. Corresponding chimeric DNA fragments were loaded as  
 concentration standards on agarose gels or labeled with [<sup>32</sup>P]-dCTP and used as probes  
 20 for Southern analysis. The number of viral genomes per DNA sample was calculated  
 after quantitative Phospho-imager analysis. In preliminary experiments, no preferential  
 hybridization of chimeric DNA probes to DNA of any particular viral serotype was  
 detected.

## 25 RESULTS

### Replication of selected serotypes in K562 and CD34+ cells:

Adsorption/internalization studies do not ultimately prove viral transduction, a process  
 30 often defined as gene transfer that allows for viral or heterologous gene expression in  
 host cells. Intracellular trafficking, including endosomal lysis, transport to the nucleus,

and nuclear import of the viral genome, depends on structural capsid proteins and thus, varies between different serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065). We believed that analysis of viral gene expression would be a means to verify successful nuclear import of viral genomes and that this would be a good criterion for selection of serotype/s able to efficiently infect our target cells. To do this, we used a protocol, which allows for the detection of Ad replication in infected cells. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. We utilized a site-specific methylation strategy to monitor viral DNA replication within infected cells (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). Methylated Ad serotypes were produced by the addition of a methyl group onto the N6 position of the adenine base of Xho I sites, CTCGAG, during propagation of the viruses in 293 cells expressing the Xho I isoschizomer PaeR 7 methyltransferase (PMT) (Kwoh, T. J., et al. 1986. *Proc.Natl.Acad.Sci. USA*. 83:7713-7717) (293 PMTcells). Loss of methylation through viral replication restores Xho I cleavage and can be detected by Southern blots of Xho I-digested genomic DNA from infected cells.

Ad replication studies were performed in K562 and CD34+ cells with Ad9 and Ad35, in comparison to Ad5. For replication studies, the infectious titer (in pfu/ml) and genome titer (in genomes per ml) were determined (by plaque assay on 293 cells or by quantitative Southern blot, respectively) for methylated and unmethylated Ad5, Ad9, and Ad35 (Table 2). The ratio of pfu to genome titer was comparable for methylated and unmethylated virus demonstrating that DNA methylation had not altered transduction properties. About 85% of (Ad5, 9, and 35) virus used for infection was methylated as calculated based on the intensity of fragments specific for methylated and non-methylated viral DNA present in the viral load (Fig. 15). The numbers of genomes detected after adsorption (1 hour, 0°C) or internalization (2 hours 37°C) correlated well with studies shown in Fig. 12. Ad9 and Ad35 interacted more efficiently than Ad5 with K562 and CD34+ cells. Dose-dependent replication studies in K562 and CD34+ cells were performed with the same genome numbers of Ad5, 9, and 35 (Fig. 15). The replication rate was measured based on the ratio of methylated to demethylated viral

DNA after infection with different MOIs (2100, 420, and 105 genomes per cell). In K562 cells, efficient replication (100% conversion from methylated to unmethylated DNA) was detected for Ad5 at MOI  $\geq 2100$ , for Ad9 at MOI  $\geq 420$ , and for Ad35 at MOI  $\geq 105$ . This demonstrated that Ad35 transduced K562 cells with the highest efficiency.

5 In CD34+ cells, the replication rate was 100% for Ad5 and 31% for Ad9 after infection with MOI 420. Although methylated Ad35 viral DNA was present in CD34+ cells, viral replication was undetectable for Ad35. In summary, while viral replication studies in K562 cells confirmed data obtained for Ad5, 9, and 35 adsorption and internalization, there was a discrepancy between earlier results and the poor replication of Ad9 and,

10 particularly, Ad35 in CD34+ cells. As outlined later, replication analysis in heterogeneous cell populations, like CD34+ cells, may not allow for definitive conclusions on tropism of a particular serotype.

Taking all the screening data together, Ad9 and Ad35 emerged as the variants with the

15 strongest tropism for K562 and CD34+ cells. It is thought that Ad9 can bind to CAR, however, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology*. 70:7614-7621). This entry strategy may be not optimal for efficient infection of CD34+ cells as only less than 17% of them express  $\alpha_v$ -integrins (Fig. 10). Therefore, we decided to concentrate on Ad35 as a source for heterologous fiber to be used for

20 construction of a chimeric vector based on an Ad5 backbone.

**TABLE IV**

**Results from the infectivity assay which determines the optical particle-to-PFU (OPU/PFU) ratio using 293 cells**

25

| Virus | OPU (A260)            | PFU                   | OPU/PFU ratio |
|-------|-----------------------|-----------------------|---------------|
| Ad5   | $1.4 \times 10^{12}$  | $1.06 \times 10^{11}$ | 13            |
| Ad9   | $4.61 \times 10^{11}$ | $2.6 \times 10^8$     | 1773          |

**DISCUSSION**

Viral replication studies in K562 cells confirmed the data obtained for Ad5, 9, and 35 adsorption and internalization. However, there was a discrepancy between the interaction data and the replication data in CD34+ cells where Ad9 replicated only poorly and no replication was seen for Ad35. Ad replication is only initiated upon the production of a critical threshold of early viral proteins, which in turn, is directly dependent on the number of viral genomes present in the nuclei of infected cells. Therefore, the outcome of replication studies may be affected by the rate of nuclear import of viral genomes, by the activity of viral promoters, and/or the intracellular stability of viral DNA/RNA. These parameters may vary, on one hand, between different subsets of CD34+, and/or, on the other hand, between different Ad serotypes. In conclusion, the viral replication analyses performed with different Ad serotypes in CD34+ cells may not predict the actual transduction properties of chimeric vectors based on Ad5 backbone. This implies that attempts to produce gene transfer vectors based on Ad genomes other than Ad5 should be exercised with caution.

Recently, an Ad serotype screening strategy was used to identify variants with tropism for primary fetal rat CNS cortex cells or human umbilical vein endothelial cells. The optimal serotype (Ad17) was selected based on immunohistochemistry for hexon production 48 hours after infection (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540). However, this approach is problematic because, at least in our hands, antibodies developed against Ad5 hexon did not cross-react with other serotypes. Also, hexon is expressed only after onset of replication. As outlined above, the kinetics of intracellular trafficking, viral gene expression, and replication significantly vary between serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065).

In addition to being the most efficient serotype in terms of interaction with CD34+ cells, Ad35 is also interesting because it interacts with receptor/s different from the Ad5 and Ad3. Ad35 and Ad5GFP/F35 attachment was not inhibited by Ad5 or anti-CAR antibodies suggesting that Ad35 binding was CAR independent. First, Ad5 did not compete with Ad35 and Ad5GFP/F35 during internalization and infection indicating that

$\alpha_3\beta_{3/5}$  integrins are not involved in viral entry. Second, function-blocking antibodies against  $\alpha_v$ -integrins did not compete with Ad35 and Ad5GFP/F35 for internalization into K562 cells, whereas these antibodies did inhibit Ad5 internalization. And third, in contrast to Ad5 based vectors, GFP expression after infection with Ad5GFP/F35 was not restricted to  $\alpha_v$ -integrin-expressing CD34+ cells. From these facts, we conclude that infection with Ad35 and the chimeric Ad5GFP/F35 vector does not involve  $\alpha_v$ -integrins. In this context, the presence or absence of RGD motifs within Ad35 penton base remains to be determined by sequencing the corresponding genome region. Cross-competition assays demonstrated that Ad35 and Ad5GFP/F35 bind to a receptor that is different from the Ad3 receptor. Although Ad3 and 35 belong to the same subgroup, they have been divided into two DNA homology clusters. B1 and B2; the amino acids composing their fibers are only 60% homologous. Furthermore, the target tissues for both viruses are different; Ad3 can cause acute respiratory infections, whereas Ad35 is associated with kidney infection (Horwitz, M. S. 1996. *Adenoviruses*, p. 2149-2171. In B. N. Fields, Knipe, D.M., Howley, P.M. (ed.), *Virology*, vol. 2. Lippincott-Raven Publishers Inc., Philadelphia). Therefore, it was not surprising to see that Ad3 and Ad35 recognize different receptors.

In conclusion, Ad35 and the chimeric vector enter the cells by a CAR- and  $\alpha_v$ -integrin independent pathway. We believe that Ad35 and the chimeric vector binds primarily to its fiber receptor and that this interaction is sufficient to trigger internalization. On the other hand, Ad35 internalization may involve cellular proteins other than  $\alpha_v$ -integrins. These membrane proteins can overlap with those for Ad3 internalization and represent  $\beta_2$  integrins, which protrude more from the cell surface than  $\alpha_v$ -integrins (Huang, S., et al. 1996. *J. Virology*. 70:4502-4508).

According to EM studies of negative contrast-stained adenoviral suspensions, the percentage of deficient particles for all adenoviral serotypes tested does not exceed 5%. However, plaque assays reveal that the ability to form plaques in 293 cells is significantly different for tested serotypes. The optical particle-to-PFU (OPU/PFU) ratio obtained is 13 for Ad5, which is in good agreement with the previously estimated ratio for this

adenoviral serotype. Importantly, this ratio is about three times higher for adenovirus serotype 35 and more than 150-fold higher for adenovirus serotype 9. Furthermore, quantitative Southern blot using chimeric Ad5/9 and Ad5/35 DNA probes is used to determine the ratio between the genome and transducing titer. This study confirms the data obtained by plaque assay. Quantitative replication assay of these adenoviruses in K562 and CD34+ cells also confirms the ability of Ad9 and Ad34 to more efficiently attach to these cell types. The replication of viral genomes is observed for Ad9 and Ad34 at lower MOIs of infection, compared to Ad5. In conclusion, the data obtained for different serotypes in attachment and internalization are in good agreement with the infectivity data in target cells.

**E. Attachment and internalization of different adenoviral serotypes into primary dendritic cells, JAWSII, MCF-7 and REVC cells.**

As a proof of principle, the serotype screening strategy can be employed for other important target cells which are refractory to Ad5 infection.

***RESULTS***

RECV cells are endothelial cells which have to be targeted for approaches that are aimed to gene therapy of restenosis, atherosclerosis, inflammation etc. MCF-7 cells are breast cancer cells isolated from liver metastases which are important targets for tumor gene therapy. The human adenovirus serotypes 3, 5, 9, 35 and 41 are tested to see whether they can attach to and can be internalized by mouse primary dendritic cells, JAWSII cells, MCF-7-human breast cancer cells and REVC endothelial cells. None of the adenoviral serotype tested can efficiently attach to primary dendritic cells. Adenovirus serotype 3 is able to efficiently attach to REVC endothelial cells (about 400 virus particles per cell are attached and about 300 are internalized). In comparison, only 50 Ad5 particles are able to attach to and even fewer are internalized in these REVC. The human breast cancer cells (MCF-7) are previously shown to be refractory to Ad5

infection at low MOIs. However, Ad3 and more efficiently, Ad35 attach to and internalize into MCF-7 cells.

### DISCUSSION

5

The data presented herein indicate that different human adenovirus serotypes recognize different cellular receptors and can therefore infect cell types that are refractory to Ad5 infection. There are adenoviral serotypes that can more efficiently attach and internalize than Ad5 for human CD34+ cells, REVC, K562 and MCF-7 cells. This finding provides  
10 a basis for the construction of chimeric adenoviral vectors which are Ad5 vectors containing receptor ligands derived from other serotypes.

#### **F. Infection studies on primary human bone marrow cells.**

15 Since established erythroleukemic cell lines do not represent an adequate model for the ultimate hematopoietic stem cell that has to be targeted in patients in order to achieve long-term reconstitution with genetically modified cells, normal primary human bone marrow cells are used for the initial infection/retargeting studies.

### 20 RESULTS

In a first set of tropism studies with different Ad serotypes, whole bone marrow cell suspensions can be used without preselection. This is advantageous because the tropism of various adenovirus serotypes or genetically retargeted vectors can be analyzed on a  
25 broad spectrum of progenitor subpopulations representing myeloid, erythroid, megakaryocytic, lymphoid, dendritic, and monocytic lineages. For short term (< 5 hours) infection studies, bone marrow suspensions can be cultured in IMDM supplemented with 10% FCS,  $\beta$ -mercaptoethanol, and 10u/ml IL-3 for ensuring cell viability.

30

Mononucleated cell assays:

Mononucleated bone marrow cells can be incubated with MOI 1, 10, 100, or 1000 pfu/cell of the various adenovirus types for a short time. Paraffin sections or cytopspins of infected bone marrow cells can be analyzed for nuclear-localized, labeled viral DNA. BrdU labeling can be visualized by immunofluorescence with anti-BrdU antibodies; <sup>32</sup>P-tagged viral DNA can be detected by incubation with photo-emulsion. In addition, the same cell material can be analyzed for morphology after specific histo-staining (e.g. Wright-, Hemo3 staining). If required, commercially available antibodies can be used to specific cell surface markers conjugated directly to different fluorochromes (FITC (green), TRIT., RPE, (red), RPE-Cy5, AMCA (blue)) to completely characterize infected bone marrow subpopulations. Colocalization of BrdU-labeled viral DNA (e.g. as FITC signal) with membrane markers signifying infection of specific cell types can be demonstrated; for example, potential stem cells/early progenitors (CD34<sup>+</sup>, CD38<sup>-</sup>), megakaryocytes (CD41a<sup>+</sup>), erythroid cells (glycophorin A<sup>+</sup>), dendritic cells (CD11a<sup>+</sup>), monocytes (CD14<sup>+</sup>), or myeloid cells (CD15<sup>+</sup>), etc. The morphological analysis of infected bone marrow subsets gives a first information whether specific adenovirus serotypes can target primitive cell types.

20 *DISCUSSION*

Since the different wild-type adenoviruses do not express a uniform marker gene and do not integrate and since detection of tagged viral DNA cannot be done on live cells, it is not possible, at this point, to characterize infected cells for clonogenic or repopulation capacities. Therefore, adenovirus serotypes for retargeting studies are selected, based on their ability to infect in vitro purified CD34<sup>+</sup> cells at low MOIs. This subset of bone marrow cells is known to contain long-term reconstituting cells. Infection studies with different adenovirus serotypes can be repeated on purified CD34<sup>+</sup> cells (cultured in IMDM +10% FCS,  $\beta$ -mercaptoethanol, and 10 units/ml IL-3) as described above. Purification of CD34<sup>+</sup> cells can be performed by direct immunoadherence on anti-CD34 monoclonal antibody-coated plates or on MiniMacs columns as described by



Papayannopoulou (Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69; Papayannopoulou, T. et al., 1993, *Blood*, 81:229). The purity of isolated CD34+ cells ranges routinely from 80-95%. Analogous infection studies can be repeated with selected adenovirus types on CD34+/CD38- subsets.

5

To confirm productive infection purified CD34+ cells can be infected with selected (methylase-tagged) serotypes and analyze viral DNA replication. Cultures of purified human bone marrow CD34+ cells can be used for the transduction and integration studies as a model for HSCs.

10

It was recently demonstrated that HSC activity does exist in CD34-negative human bone marrow subsets (Bathia, M. et al., 1998, *Nature Medicine*, 4:1038-45; Osawa, M., et al., 1996, *Science*, 273:242-5; Goodell, M. et al., 1997, *Nature Medicine*, 3:1337-45; Zanjani, E. D. et al., 1998, *Exp. Hematology*, 26:353-60). Lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup> cells can be tested in the retargeting and transduction studies in combination with repopulation assays in SCID-NOD mice.

15

#### G. Cloning and insertion of the fiber gene.

#### 20 *METHODS*

##### PCR-cloning of the corresponding fiber gene and insertion into Ad5 based shuttle plasmids instead of the endogenous AD5 fiber:

25 One or several adenoviruses with tropism to CD34+ or other HSC containing population is selected for further studies described herein. The complete coding region for fiber varies between 1-2kb, depending on the virus type. The fiber encoding sequences can be obtained by PCR with Pfu polymerase from viral DNA isolated from purified particles of the selected virus types. The corresponding primers can be designed based on the fiber sequences available from the EMBL gene bank. The PCR products are cloned as PacI-Ball fragment into pCD4 (Figure 10), a shuttle vector for recombination of RecA+ E.

30

coli. In pCD4, the heterologous fiber gene is flanked on both sides with Ad5 sequences, which are homologous to regions directly adjacent to the fiber reading frame in Ad5. As an Ad5 (shuttle vector) derived template for recombination, pCD1, a pBHG 10 (Microbix, Toronto, Canada) derivative can be used. The recombination procedure is performed according to a protocol routinely used for recombinant adenovirus generation (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Routinely, 90% of the resulting plasmids are accurately recombined. The junctions between the heterologous fiber (X) and Ad5 sequences can be sequenced to confirm the accuracy of recombination. The resulting plasmid is named pAd5fiberX (pAd5<sup>fx</sup>). The resulting product is used to generate pAd5<sup>fx</sup>-based Ad.AAV containing the heterologous fiber gene.

Construction of chimeric Ad vectors:

For transduction studies, two Ad vectors were constructed: Ad5GFP and Ad5GFP/F35, containing a chimeric Ad5/35 fiber gene. Both adenoviral vectors contained a 2.3kb, CMV promoter driven EGFP gene [derived from pEGFP-1, (Clontech, Palo Alto, CA)] inserted into the E3 region of Ad5. The EGFP expression cassette was cloned between Ad5 sequences 25,191-28,191 and 30,818-32,507 into a shuttle plasmid, which contained the E3 deletion described for pBHG10 (Microbix, Toronto, Canada). The resulting plasmid was named pAdGFP. For the chimeric vector, the Ad5 fiber gene in pAdGFP was substituted by an Ad5/35 chimeric fiber gene generated by the two-step PCR protocol outlined above. In the first PCR step, three DNA fragments corresponding to i) the Ad5 fiber 5'-nontranslated region and the first 132 bp of the fiber tail domain (nt 30,818-31,174), ii) the Ad35 shaft and knob domains (nt 132-991), and iii) the Ad5 E4 region including the Ad5 fiber polyadenylation signal (nt 32,775-33,651) were amplified by Pfu-Turbo DNA polymerase. The following primers were used: for the Ad5 tail, Ad5F-2 (nt 30,798-30,825) 5'-CGC GAT ATC GAT TGG ATC CAT TAA CTA-3' and Ad5R-2 (nt 31,174-31,153) 5'-CAG GGG GAC TCT CTT GAA ACC CAT T-3'; for the Ad35 shaft and knob, primers Ad5/35F and Ad5/35R (see above); for the Ad5E4 and polyA, primers Ad5F-1 and Ad5R-1 (see above). After 10 PCR cycles, the products were purified by agarose gel electrophoresis, combined, and then subjected to a second PCR

with primers Ad5F-2 and Ad5R-1. The resulting 2115 bp-long chimeric fiber gene contained the Ad5 tail and the Ad35 shaft and knob domains. This product was used as a substitute for the Sall/XbaI Ad5 fiber gene containing fragment in pAdGFP. The resulting plasmid was named pAdGFP/F35. To generate full-length E1/E3 vector genomes, pAdGFP and pAdGFP/F35 were inserted in pAdHM4 (Mizuguchi, H., Kay, M.A. 1998. *Human Gene Therapy*. 9:2577-2583) by recombination in *E.coli* (Chartier, C., E. et al. 1996. *Journal of Virology*. 70:4805-4810). To do this, the RecA+ *E.coli* strain BJ5183 was co-transformed with pAdHM4 linearized by SrfI mixed with the XbaI fragments containing the GFP genes, the Ad5 or Ad5/35 fiber genes, and the Ad5 homology regions. The resulting recombinants were analyzed by restriction analysis. Correct recombinants were amplified in *E.coli* HB101 and purified by double CsCl gradient banding. The plasmids were named pAd5GFP and pAd5GFP/F35. The correct structure of the Ad5/35 chimeric fiber gene was confirmed by endonuclease digestion and sequencing part of pAd5GFP/F35. To produce the corresponding viruses, pAd5GFP and pAd5GFP/F35 were digested with PacI to release the viral genomes and transfected onto 293 cells as described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). Plaques developed 7 to 10 days post-transfection in overlaid cultures. Recombinant viruses were propagated in 293 cells and purified by standard methods described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

#### Hemagglutination assay:

Twenty-five microliters of serial dilutions of Ad5, Ad35, or chimeric Ad5GFP/F35 virions in McIlvaine-NaCl buffer (0.1 M citric acid, 0.2 M Na<sub>2</sub>HPO<sub>4</sub> [pH 7.2], diluted 1:50 with 0.87% NaCl) were loaded onto 96 well plates. To each dilution, 25 µl of a 1% suspension of monkey erythrocytes (in McIlvaine-NaCl buffer) was added. The sedimentation pattern was determined after incubation for 1 hour at 37°C. All tests were performed in quadruplicates in at least two independent experiments.

Southern blot:

Extraction of genomic DNA, labeling of DNA fragments and hybridization were performed as described earlier (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

**RESULTS**Construction/Characterization of chimeric fiber:

Previously, it has been shown that exchanging the fiber knob was sufficient to alter the tropism of chimeric Ad vectors (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V., et al. 1998. *J. Virology*. 72:1844-1852; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). As outlined above, the length of the fiber shaft may critically determine the entry strategy of a particular serotype. Therefore, we decided to replace not only the Ad5 fiber knob but also the shaft. The chimeric Ad5/35 fiber contained the Ad5 tail (amino acid: 1-44) necessary for interaction with the Ad5 penton base linked to 279 amino acids from Ad35 including the shaft with 7  $\beta$ -sheets and the knob (Fig. 16A). The endogenous Ad5 fiber polyA signal was used to terminate transcription of the chimeric fiber RNA. The combination of the Ad5 capsid including the RGD motif containing penton base with a short-shafted fiber could be risky because the natural distance between the fiber knob and the RGD motifs was disturbed. The Ad5 fiber was substituted by the chimeric fiber sequences based on an E1/E3 deleted Ad vector. This vector carried a CMV promoter-GFP reporter gene cassette inserted into the E3 region. The corresponding chimeric virus (Ad5GFP/F35) was produced in 293 cells at a titer of  $>2 \times 10^{12}$  genomes per ml. For comparison, an E1/E3 deleted Ad vector containing the original Ad5 fiber gene and the GFP expression cassette was generated (Ad5GFP). The titer and the ratio of physical to infectious particles was similar between Ad5GFP and Ad5GFP/F35 indicating that the fiber modification did not significantly alter the stability and/or growth properties of the chimeric vector. The correctness of the fiber modification was confirmed by restriction analysis of the Ad5GFP/F35 viral genome followed by

Southern blot hybridization (Fig. 16B), direct sequencing of the fiber-coding region, and a functional test for hemagglutination (HA) of monkey erythrocytes. The agglutination of erythrocytes is fiber knob-mediated; it is known that Ad5 does not agglutinate monkey erythrocytes whereas Ad35 efficiently does (Pring-Akerblom, P., et al. 1998, *J. Virology*, 72:2297-2304). In HA tests, Ad5GFP/F35 agglutinated monkey erythrocytes with the same efficiency as Ad35 at dilutions of up to 1:512. In contrast, no hemagglutination was observed with equivalent Ad5 dilutions. This clearly confirmed the functional activity of the chimeric Ad5/35 fiber incorporated into Ad5 capsid.

10 Generation of chimeric adenoviral vectors (Ad.AAV<sup>fx</sup>) with heterologous fiber molecules: Adenoviruses with chimeric Ad5-Ad3 fiber are viable and can be produced at high titers (Krasnykh, V., et al., 1996, *J. of Virology*, 70, 6839-6846; Stevenson, S. C. et al., 1997, *J. Virology*, 71:4782-90). In order to test whether the fiber substitution described herein affects production or stability of adenoviruses, two E1-deleted first-  
15 generation, adenoviral vectors are produced with the AAV- $\beta$ gal cassette in 293 cells using standard protocols. The vector is generated by recombination of pAd.AAV-BG (Fig. 17) with pCD1 (containing the endogenous Ad5 fiber); the other vector (with heterologous fiber) is the recombination product of pAd.AAV  $\beta$ Gal and pAd5fiberX (pAd5<sup>fx</sup>). Virus from single plaques is amplified on 293 cells. The production yield per  
20 293 cell can be determined by plaque-titering of 293 cell Lysates. It is anticipated that the fiber modification will not critically affect the stability of chimeric vectors. Finally, bone marrow cells can be infected with the retargeted vectors. Two days after infection, live-cell cytometry is performed for  $\beta$ -gal expression using as substrate Fluorescein di- $\beta$ -D-Galactopyranoside (FDG) (Cantwell, M.J. et al., 1996 *Blood* 88, 4676-4683; Neering, S. et al., 1996, *Blood*, 88:1147-55; Fiering, S. N. et al., 1991, *Cytometry*, 12:291; Mohler, W. et al., 1996, *PNAS*, 93:57) and the infected cells are characterized for morphology and surface markers. Before and during infection, bone marrow cells can be cultured in IMDM/FCS supplemented with thrombopoietin (Tpo), which supports the survival of HSC (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996,  
25 *Experimental Hematology*, 24:660-69). Alternatively, retargeted vectors can be  
30

generated with the AAV-GFP (green fluorescence protein) cassette and perform FACS analysis on transduced cells based on GFP and surface marker expression.

#### **H. Competition studies of chimeric fiber protein Ad5/35.**

5

##### **Competition studies:**

Cross-competition studies between Ad5, 35, and Ad5GFP/F35 (Fig. 18) for binding and internalization were performed in order to investigate in more detail the pathways which are used by the chimeric vector to infect target cells. Wild-type Ad35 and the chimeric vector Ad5GFP/F35 could recognize the same primary receptor as they competed with each other for the attachment to K562 cells (Fig. 19A, upper panel). This primary receptor is different from that used by Ad5, since neither Ad5 viral particles nor anti-CAR monoclonal antibodies (Fig. 19B, upper panel) were able to abrogate Ad35 or Ad5GFP/F35 binding. In competition studies for internalization, Ad35 and Ad5GFP/F35 competed with each other with equal efficiency. Ad5 and anti- $\alpha_v$ -integrin monoclonal antibodies (L230) (Figs. 19C, D; lower panel) did not inhibit internalization of Ad35 or the chimeric virus. To consolidate this data, K562 cells were infected with Ad5GFP and Ad5GFP/F35 after prior incubation of cells with anti-CAR or anti- $\alpha_v$ -integrins monoclonal antibodies followed by analysis of GFP-expressing cells. The transduction data mirror the results obtained in adsorption/internalization studies. In summary, this demonstrated that Ad35 and Ad5GFP/F35 use a CAR and  $\alpha_v$ -integrin-independent pathway for infection of K562 cells; the structural elements which account for these specific properties are located within the Ad35 fiber and can be transplanted into Ad5 by fiber substitution.

Ad3 can efficiently interact with K562 cells (Fig. 12), although Ad3 and Ad35 belong to the same subgroup (B), the homology between amino acid sequences of their fibers is only about 60%. Therefore, we decided to test whether Ad3 could compete with Ad35 and Ad5GFP/F35 for attachment and internalization (Fig. 20). These studies demonstrated that Ad35 binding was not inhibited by Ad3 indicating the use of different

30

receptors. Interestingly, Ad3 slightly inhibited attachment of Ad5GFP/F35 (Fig. 20A, left panel). In addition to binding to the receptor common for the Ad35 and Ad5GFP/F35 fiber, the chimeric capsid (e.g. the Ad5 penton RGD motifs) may also interact with a second cellular receptor that overlaps with elements involved in Ad3 binding. In cross-competition for internalization, pre-incubation of cells at 37°C with Ad35 and with chimeric virus significantly decreased internalization of [<sup>3</sup>H]-labeled Ad3 (Fig. 20D, right panel). In the reverse experiment, Ad3 as competitor decreased the level of internalization by 30% for both, Ad35 and the chimeric virus (Fig. 20B, right panel). As expected, Ad5 and Ad3 did not compete for adsorption or internalization. As shown before (Fig. 19B), anti-CAR and anti- $\alpha_v$ -integrin antibodies did not block Ad3 interaction with K652 cells. In summary, we concluded that Ad35 and Ad5GFP/F35 bind to receptor/s different from that of Ad3, although they can use common structural elements for internalization, which are different from  $\alpha_v$ -integrins.

15 Infection studies with chimeric virus:

It is established that Ad5GFP/F35 infected K562 cells by a CAR and  $\alpha_v$ -independent pathway. It is possible that this property allows for efficient transduction of non-cycling CD34+ cells, which express scarcely CAR and  $\alpha_v$ -integrins. To test this, the transduction properties of Ad5GFP and Ad5GFP/F35 vectors were analyzed on CD34+ cells, K562, and HeLa cells. Fig. 21 shows the percentage of transduced, GFP expressing cells depending on the MOI used for infection. Nearly 100% of HeLa cells were transduced with Ad5GFP and Ad5GFP/F35 at MOIs of  $\geq 25$ . More than 95% of the K562 cells were transduced with Ad5GFP/F35 at MOIs of  $\geq 100$ , whereas the transduction rate was significantly lower with Ad5 where it increased with the MOI reaching a plateau at ~70% GFP-positive cells after infection with an MOI of 400. Transduction of CD34+ cells was about three fold more efficient with Ad5GFP/F35 than with Ad5GFP at all MOIs analyzed. Interestingly, at higher MOIs, the transduction rate did not rise proportionally with the viral dose and soon reached a plateau indicating that in both cases only specific subset/s of CD34+ cells were permissive to infection. In order to characterize in more detail these specific, permissive subset/s, additional transduction

studies were performed. First, the percentage of GFP expressing cells was determined in CD34+ fractions that were stained for  $\alpha_v$ -integrins or CARs (Fig. 22). The low number of CAR positive CD34+ cells complicated accurate co-labeling studies, and there was no correlation between CAR expression and the proportion of transduced cells among CD34+ cells infected with Ad5GFP or Ad5GFP/F35. Interestingly, for Ad5GFP, 65% of all GFP expressing cells were positive for  $\alpha_v$ -integrins, whereas less than 22% of GFP positive cells infected with the chimeric virus stained positive for  $\alpha_v$ -integrin expression. While only 17% of the whole CD34+ population expressed GFP after Ad5GFP infection, the percentage of GFP-expressing cells in the CD34+/ $\alpha_v$ -integrins positive fraction was 50%. This indicates that Ad5GFP vector-mediated GFP expression was preferentially localized to  $\alpha_v$ -integrin positive CD34+ subsets, whereas after infection with the Ad5GFP/F35 vector, GFP was expressed in a broader spectrum of CD34+ cells with most of them being  $\alpha_v$ -integrin-negative.

Next, transduced cells were simultaneously analyzed for GFP as well as for CD34 and CD117 markers. As mentioned before, only about 90% of all cells used in our analysis were positive for CD34 at the time of infection, hence the multiparameter analysis for CD34 and GFP. A population of CD34+ cells is extraordinarily heterogeneous in morphology and stem cell capacity. The subpopulation of CD34+ and CD117+ cells resembles very primitive hematopoietic cells (*Ikuta, K., Weissman, I.L. 1992 Proc. Natl. Acad. Sci. USA. 89:1502-1506; Simmons, P. J., et al. 1994. Expl. Hematology. 22:157-165*). Fig. 23 summarizes the analyses of GFP expression in correlation with these specific stem cell markers. While 54% of cells infected with chimeric vector were positive for GFP and CD34+, only 25% of cells infected with Ad5GFP expressed the transgene and CD34+ marker (Fig. 23A, lower panel). More importantly, based on GFP expression, the chimeric virus transduced 80% of c-kit positive cells, whereas the Ad5-based vector transduced only 36% (Fig. 23A, middle panel). In an additional experiment, CD34+ cells were sorted for CD117 expression prior to infection with Ad5GFP or Ad5GFP/F35 and, 24 hours post-infection, GFP expression was analyzed in this specific fraction (Fig. 23B). This analysis revealed that the chimeric vectors transduced 4 fold more CD34+/CD117+ than the Ad5GFP vector.



In conclusion, these results demonstrated that the chimeric Ad5GFP/F35 vector was clearly superior to the Ad5GFP vector in targeting and transduction of CD34+ cells. Furthermore, the data suggest that the spectrum of CD34+ cell subsets permissive for Ad  
5 infection was significantly different for the chimeric vector than for the Ad5 vector.

Analysis of viral genomes within CD34+ cells infected with the Ad5 and chimeric vectors:

10 So far, the transduction rate of CD34+ cells was measured based on GFP expression after infection with Ad5GFP and Ad5GFP/F35. Considering the extraordinary heterogeneity of CD34+ cells in morphological and functional parameters, GFP may not be expressed in all cell types that were efficiently infected. Reasons for this include that the CMV promoter may not be active in all cell types or that the regulation of transgene expression  
15 could differ between subsets on a post-transcriptional or post-translational level. To test this, we quantified the number of intracellular (transduced) viral genomes within GFP positive and GFP negative fractions of CD34+ cells infected with Ad5GFP and Ad5GFP/F35. To do this, twenty-four hours after infection, CD34+ cells were sorted for GFP positive and GFP negative fractions, which were subsequently used to isolate  
20 genomic DNA together with transduced viral DNA. The number of viral genomes was determined by quantitative Southern blot as described for Fig. 15. Per GFP-positive CD34+ cell, about 270 copies of the Ad5GFP/F35 viral genome were detected. Interestingly, a remarkable 200 copies of the Ad5GFP/F35 viral genome were found per GFP-negative CD34+ cell (Fig. 24A and 25). This demonstrated that not all infected cells  
25 expressed GFP and implies that the actual transduction rate was higher than 54% (GFP-positive cells). We concluded that the CMV promoter was not active in all transduced CD34+ subsets. No Ad5GFP vector specific signal was detected within infected CD34+ (GFP positive or negative) fractions by Southern blot which had a detection limit of 14 viral genomes per cell. From this, we can conclude that the vector DNA concentration per  
30 transduced cell was at least 20 times higher for Ad5GFP/F35 than for Ad5GFP.

Ad5GFP DNA was only detectable in DNA samples from infected CD34+ cells by Southern blot after prior PCR amplification with vector specific primers (Fig. 24B and 25). This indicates that the replication deficient Ad5 vector is present but at a very low copy number, which may be limited by intracellular genome stability. Using the PCR-Southern detection method, Ad5 vector DNA was also detected in GFP negative cells, supporting that the CMV promoter may not have been the optimal choice for transduction studies. It is notable that studies by others on viral genome analyses after infection of CD34+ cells with Ad5 vectors were performed only after prior PCR amplification (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Neering, S. J., et al. 1996.. *Blood*. 88:1147-1155).

### DISCUSSION

The chimeric Ad5GFP/F35 vector has binding and internalization properties similar to Ad35. Therefore, the fiber substitution was sufficient to swap cell tropism from Ad5 to Ad35. The Ad5GFP/F35 capsid chimera contained the short-shafted Ad35 fiber incorporated into an Ad5 capsid, instead of the naturally occurring long-shafted Ad5 fiber. During Ad5 infection, interaction between the penton base and integrins is required to induce viral internalization. For this interaction, the length of fiber shaft and the precise spatial arrangement of knob and RGD motifs are critical for the virus entry strategy. The natural spatial arrangement is disturbed when short-shafted heterologous fibers are inserted into the Ad5 capsid. Interestingly, the Ad5/35 capsid chimera allows for efficient infection, suggesting that the protruding RGD motives in the Ad5 penton base do not affect the interaction with the primary Ad35 receptor. So far, most of the chimeric viruses were generated by substituting only the Ad5 knob while maintaining the long Ad5 fiber shaft (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V. N., et al. 1996. *J. Virology*. 70:6839-6846; Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). The exception was an Ad5/7 chimeric virus (Gall, J., et al. 1996. *J. Virology*. 70:2116-2123), where the whole Ad5 fiber was substituted by the short-shafted Ad7 fiber. However, similar to the parental Ad5, the Ad5/7 chimera still required  $\alpha_v$ -integrins for infection.

This Ad5GFP/F35 chimera is the first demonstration that despite the presence of RGD motifs within the Ad5 penton, the chimeric virus uses cell entry pathways determined primarily by the receptor specificity of the short-shafted heterologous fiber. This does not  
5 exclude that interaction with a secondary receptor may increase binding affinity. The latter is supported by the observation that Ad35 and Ad5GFP/F35 slightly differed in their ability to compete with Ad5 or Ad3 for binding. It is possible that Ad5/35 attachment involves, in addition to the high affinity fiber binding, interaction between Ad5 capsid proteins (e.g. RGD motifs) and secondary receptor/s that overlap with those  
10 used by Ad3 and Ad5.

This data indicate that infection with Ad5-based vectors is restricted to a specific subset of CD34+ cells. The percentage of GFP expressing cells after Ad5GFP infection of CD34+ cells reached a plateau at MOIs higher than 100 indicating that only a limited  
15 fraction of CD34+ cells was permissive to Ad5. Also, strong replication of wild type Ad5 in infected CD34+ cells may be the result of preferential transduction of a specific subpopulation of CD34+ resulting in a expression of early viral genes at a level sufficient to initiate viral replication. The presence of a specific subpopulation of CD34+ cells permissive to Ad5-vector infection was suggested by others (Byk, T., et al. 1998. *Human*  
20 *Gene Therapy*. 9:2493-2502; Neering, S. J., et al. 1996.. *Blood*. 88:1147-1155). In the present report, we further characterized this subpopulation and demonstrated that Ad5-based vectors preferentially infected  $\alpha_v$ -integrin positive CD34+ cells. Integrins (including  $\alpha_v$ ) are thought to be important for homing and trafficking of transplanted hematopoietic cells, however little is known about the correlation between  $\alpha_v$ -integrin  
25 expression and the differentiation status of hematopoietic cells (Papayannopoulou, T., Craddock, C. 1997. *Acta Haematol.* 97:97-104; Roy, V., Verfaillie, C.M. 1999. *Exp. Hematol.* 27:302-312). There was no clear correlation between CAR and GFP expression suggesting that Ad5GFP may be able to use another membrane protein as a primary receptor. Alternatively, Ad5GFP transduction observed at an MOI of 200-400 could be  
30 the result of direct interaction between virus and  $\alpha_v$ -integrins triggering internalization, which may be the preferred pathway in the absence of CAR (Legrand, V., et al. 1999. *J.*

*Virology*. 73:907-919). Importantly, infection with the chimeric Ad5GFP/F35 vector was not restricted to the  $\alpha_v$ -positive CD34+ subpopulation.

Among CD34+ cells, the subpopulation of CD34+ and CD117+ cells resembles very  
5 primitive hematopoietic cells (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*.  
89:1502-1506; Simmons, P. J., et al. 1994. *Expl. Hematology*. 22:157-165). The receptor  
for stem cell factor, CD117 (c-kit) belongs to a tyrosine kinase family. It was previously  
shown that c-kit+, CD34+ cord blood cells contain a high fraction (16%) of  
hematopoietic progenitors (Neu, S., et al. 1996. *Leukemia Research*. 20:960-971). Early  
10 in ontogeny 34+/CD117+ cells have long-term repopulating activity (Sanchez, M. J., et  
al. 1996. *Immunity*. 5:513-525). An average of 50-60% of CD34+ cells are reported to be  
CD117 positive (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*. 89:1502-  
1506; Neu, S., et al. 1996. *Leukemia Research*. 20:960-971; Simmons, P. J., et al. 1994.  
*Expl. Hematology*. 22:157-165). In our studies, the chimeric vector expressed GFP in  
15 54% CD34+ cells and 80% of CD34+/c-kit+ cells. The actual viral transduction rate  
could be even higher because transduced Ad5GFP/F35 vector DNA was also found in  
GFP-negative fractions of infected cells. This indicates that the CMV promoter used to  
drive GFP expression in our vectors was not active in all transduced cells. We selected  
the CMV promoter for transgene expression based on published data demonstrating that  
20 PGK and CMV promoters allowed for efficient transgene expression in CD34 cells  
whereas the HTLV-I and RSV promoter were almost inactive (Byk, T., et al. 1998.  
*Human Gene Therapy*. 9:2493-2502; Case, S. S., et al. 1999. *Proc. Natl. Acad. Sci. USA*.  
96:2988-2993). On the other hand, studies by Watanabe et al. (Watanabe, T., et al. 1996.  
*Blood*. 87:5032-5039) suggest that the CMV promoter is not active or rapidly silenced in  
25 certain CD34+ subsets. Our data underscore this observation. Considering retroviral  
transduction studies, the retroviral MLV promoter may have been a better candidate for  
transduction studies in hematopoietic cells (Bregni, M., et al. 1998. *Gene Therapy*. 5:465-  
472).

30 After having demonstrated that the Ad5GFP/F35 vector efficiently transduced cells  
carrying stem cell specific markers, the next logical step would be to perform colony

assays with pre-sorted GFP positive/negative cells. However, this assay is complicated by the fact that infection with first generation Ad vectors is cytotoxic and affects the formation and growth of progenitor colonies in MC-cultures (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Watanabe, T., et al. 1996. *Blood*. 87:5032-5039). This side effect is caused by the expression of Ad proteins within transduced cells (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960; Schiedner, G., et al. 1998. *Nature Genetics*. 18:180-183; Yang, Y., et al. 1994. *Proc. Natl. Acad. Sci. USA*. 91:4407-4411). Some of these proteins (e.g E4-orf4, pTP, or E3-11.6k) have pro-apoptotic activity (Langer, S. J., Schaak, J. 1996. *Virology*. 221:172-179; Lieber, A., et al. 1998. *J. Virology*. 72:9267-9277; Shtrichman, R., Kleinberger, T. 1998. *J. Virology*. 72:2975-2983; Tollefson, A. E., A et al. 1996 *J. Virology*. 70:2296-2306). Clearly, this would affect the outcome of transduction studies with Ad5GFP/F35, which allows for the efficient transfer of viral genomes into CD34+ cells implying significant expression of viral proteins. Moreover, recently published data indicate that short-term colony assay mostly measure mature progenitors and do not represent a rigorous test for transduction of potential stem cells.

A definitive demonstration that Ad5GFP/F35 based vectors can transduce HSC requires colony assays or preferably, repopulation assays in SCID-NOD mice. We can perform these studies with gutless vectors (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) and integrating  $\square$ Ad.AAV vectors devoid of all viral genes (Lieber, A., et al. 1999. *J Virol* 73:9314-24) generated based on Ad5GFP/F35 chimeric capsids. Alternatively, gutless, retargeted vectors could be used to transiently express a retroviral receptor on CD34+ cells to increase their susceptibility to infection with retroviral vectors based on an approach that we have published earlier (Lieber, A., et al.. 1995. *Human Gene Therapy*. 6:5-11).

Our finding that Ad5GFP/F35 can efficiently transduce hematopoietic cells with potential stem cell capacity represents an important step towards stable gene transfer into HSCs and gene therapy of blood disorders. Furthermore, the virological aspects of this invention contribute to a better understanding of adenovirus cell interactions.

**I: Retargeting of Ad5 based vectors with modified fibers carrying specific ligand peptides for HSC and other cell types**

5 Another alternative to make Ad5-capsid-based vectors suitable for HSC gene therapy is to incorporate the coding sequence for HSC specific peptides into the H1 loop region of the Ad5 fiber gene. The modification of the H1-loop was successfully exercised by Krasnykh et al. with a 7 amino-acid long FLAG peptide (DYDDDDK). Using phage display peptide libraries (Pascqualini, R. et al., 1996, *Nature*, 380:364-66), Renata  
10 Pasqualini (La Jolla Cancer Research Center) reported recently, at the First Meeting of the American Society for Gene Therapy, the identification of small peptide ligands specific for bone marrow cells. The corresponding sequences encoding these peptides can be added to modify the H1 loop sequence employing site-directed mutagenesis. Optimally, the ligands should allow for the efficient internalization of adenoviral particles  
15 based on a CAR- and integrin independent pathway. Modified adenoviral vectors containing the AAVBG cassette can be produced and tested for HSC tropism as described above.

**Adenovirus peptide display:**

20

In order to retarget adenoviruses to any cell type of interest, a strategy is provided which involves creating a library of adenoviruses displaying random peptides in their fiber knobs as ligands and screening this library for adenovirus variants with tropism to a particular cell type in vitro and potentially in vivo.

25

The development of the adenovirus peptide display technique is based on the following ideas. (i) Although the tertiary structure of the Ad5 fiber knob is known, it remains unclear which domains are involved in receptor binding. There are data suggesting that receptor-binding domains partially overlap with hemagglutination domains, which are  
30 well characterized for a number of serotypes. Therefore, three intramolecular loop regions representing potential receptor binding sites can be substituted by random peptide

libraries. Eight amino acid residues in the center of the FG, or GH loops can be substituted by octameric random peptides (Figs. 26 and 27). These substitutions will replace CAR tropism and allow for infection of refractory cell types. (ii) To synthesize the oligonucleotides encoding the peptide library a novel technique to assemble pre-synthesized trinucleotides representing the codons for all 20 amino acids is employed. This avoids termination codons and assures optimal codon usage and translation in human cells. Synthesis of a completely randomized library is possible with all 20 amino acids being incorporated with the same probability and a partially randomized library with only three (in average) random amino acids substitutions per octamer at random positions with a random amino acid to maintain certain critical features of the tertiary knob structure while introducing variability. The last model is based on the distribution of amino acids present in the hypervariable CDR 1 or 2 region of immunoglobulins. (iii) To maintain a representative library size of about  $10^{10}$  different octamers per modified loop, a new cloning strategy is employed to allow for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. This strategy is based on a "seamless" cloning technique available from Stratagen. (iv) In order to produce the library of viruses, viral genomic DNA containing the modified fiber sequences is transfected into 293 cells without reduction of the library size. This critical step is done by conjugating the viral library DNA to carrier Ad5-based adenovirus via poly-lysine to assure 100% transfection efficiency. This technique allows for coupling of  $\sim 1\mu\text{g}$  of plasmid DNA (or  $\sim 1 \times 10^{10}$  adenoviral genomes) to  $10^{10}$  viral particles which can be used to infect 293/cre cells at an MOI of 10-100. Importantly, the carrier adenoviral genome has the packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express cre recombinase (293/cre). This helper virus system is routinely used to produce so-called gutless adenoviruses. Therefore, the virus progeny represents library genomes packaged into capsids containing preferentially Ad5 fibers. This is important for the next infection step into 293 cells at a MOI of 1 to assure a homogeneous fiber population on the capsid where the fibers are encoded by the packaged genome.

#### **J. Production Of Adenovirus Vectors With Increased Tropism To Hepatocytes**

An example of a G-H loop substitution to target Ad5 to hepacytes was successful. Preliminary tests demonstrated that two evolutionarily conserved regions within the malaria circumsporozoite surface protein (CS) termed RI and RII+ mediate specific  
5 interaction with hepatocytes but not with other organs (including spleen, lung, heart and brain), nor with Kupffer cells, liver endothelial cells or with other regions of the hepatocyte membrane (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Shakibaei, M. and U. Frevert, 1996, *J. Exp. Med.*, 184:1699-711). These regions are conserved among different species including *Plasmodium berghei*, *P. cynomogli*, and *P. falciparum* that  
10 infect mouse, monkey and human hepatocytes, respectively (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). Peptides derived from RI (KLKQPG) or RII (EWSPCSVTGNGIQVRIK) blocked CS binding to hepatocytes and infection by sporozoites in vivo ((Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). RI and RII+ peptides were  
15 separately inserted into Ad5-fiber knob (H-I and G-H loop) containing mutation with abolished binding to CAR and alpha-v integrins (Kirby, L. et al., 2000, *J. Virol.*, 74:2804-13; Wickham, T. J. et al., 1995, *Gene Ther.*, 2:750-6). Based on preliminary data, a short-shafted fiber was used so that the virus entry strategy predominantly depends on the interaction with the primary (hepatocyte-specific receptor). The hepatocyte-specific  
20 ligands are flanked by short glycine stretches to provide flexibility and embedded into a loop formed by two cystines. This is one of the classical strategies to incorporate ligands into a protein scaffold (Doi, N. and H. Yanagawa, 1998, *Cell Mol. Life Sci.*, 54:394-404; Koivunen, E. et al., 1995, *Biotechnology (NY)*, 13:265-70) and to guarantee their presentation at the protein surface. The biodistribution of the best variants is tested in  
25 vivo in C57B1/6 mice based on Southern blots or PCR for vector DNA in different organs. This mouse strain is known to be susceptible to infection with *P. berghei* (Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81).



**K. Production Of Adenovirus Vectors With Increased Tropism To Tumor Cells**

A similar strategy is to insert two peptides obtained after selection for tumor tropism by displaying random peptides on filamentous phages. The first double cyclic peptide  
5 (RGD-4) proved to bind specifically to integrins present on tumor vasculature (Ellerby H. M. et al., 1999, *Nat. Med.*, 5:1032-8). The second peptide targets specific matrix metalloproteinases associated with metastatic tumor cells as shown for the breast cancer cell line MDA-MB-435 (Koivunen, E. et al., 1999, *Nat. Biotechnol.*, 17:768-74). Tropism-modified vectors are tested in animal models with hepatic metastases derived  
10 from MDA-MB-435 cells (Fig. 28).

**L. Development Of A Peptide Display Technique Based On Adenoviruses**

A synthetic peptide library is described that allows adenovirus vectors to express random  
15 peptides in the G-H loop of the fiber knob domain. The technique of a phage display library is optimized to generate a library of adenoviruses displaying random peptides in their fiber knob. This library of adenovirus variants is then screened for tropism to a particular cell type in vitro and potentially in vivo. The oligonucleotides encoding the peptide library employ a novel technique to assemble pre-synthesized trinucleotides  
20 representing the codons for all 20 amino acids. This will end the termination codons and assure optimal codon usage and translation in human cells. To maintain a representative library size, a new "seamless" cloning strategy that allows for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. Transfection into 293 cells is done by  
25 conjugating the viral library DNA to carrier Ad5-based adenovirus via polylysine to assure a 100% transfection efficiency. Importantly, the carrier adenoviral genome has its packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express Cre recombinase (293/cre). The library is produced with E1-positive viruses depleted for CAR and integrin tropism. Only variant that have  
30 successfully infected the cell type of interest will replicate, resulting in de novo produced

virus. The sequence of the peptide ligand that conferred the particular tropism will then be analyzed in do novo produced virus.

### EXAMPLE III

5

#### COMBINATION NOVEL ADENOVIRAL VECTOR AND MODIFIED FIBER PROTEIN

This example describes the following studies which combine the technology of the integrating adenovirus vector that is devoid of all adenoviral genes with the modified  
10 fiber protein that retargets the vector to quiescent HSC.

##### A. Transduction studies with re-targeted vectors in HSC:

15 In order to transduce quiescent HSC and integrate into chromosomal DNA, retargeted  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors are tested for reporter gene expression, and vector integration simultaneously while analyzing their clonogenic capacity. The modified  $\Delta\text{Ad.AAV}^{\text{fx}}$  hybrid vectors contain genomes devoid of all adenoviral genes (a "gutless" adenovirus vector) packaged into Ad5 capsids with modified fibers. Rep may be incorporated into  
20 these  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors to allow for site-specific integration into AAVS1.

##### Transduction studies:

Purified human CD34+ cells in IMDM/FCS+IL-3 and SCF are infected with different  
25 doses of  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-BG}$  ( $1\text{-}10^7$  genomes per cell). CD34+ cells infected with  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-}\beta\text{Gal}$  are cultured for 2 days in suspension and sort  $\beta\text{-Gal}^+$  cells by FACS using FDG as substrate. This determines the infection efficiency.  $\beta\text{-gal}$  expressing cells are then submitted to clonogenic assays in semi-solid cultures (in two dishes per MOI) in the presence of multiple cytokines. (IL-3, SCF, Epo, G-CSF, GM-CSF, IL-7, Tpo). A  
30 first set of semi-solid cultures can be evaluated after 7 days; another set can be analyzed after 14 days. Colonies that have formed in semisolid culture can be characterized by

light microscopy and subsequently stained with X-Gal staining. Most of the vector genomes should remain episomal and can be lost with successive cell divisions. Thus, while most cells can be X-Gal positive at day 2 or day 7 after infection, most of the larger colonies (analyzed at day 14 p.i.) may not stain homogeneously for  $\beta$ -Gal. A representative number of X-Gal positive and X-Gal negative colonies can be picked and analyzed for episomal and integrated vector DNA. The outcome depends on the MOI used for infection and the integration status of the vector. These studies determine whether hybrid vectors can infect primitive progenitors.

10 Detailed characterization of hybrid vector integration:

CD34<sup>+</sup> cells can be infected with  $\Delta$ Ad.AAV<sup>fx</sup>-SNori (MOI 1-10<sup>7</sup>) and subjected to G418 selection in methyl cellulose (MC) cultures in the presence of growth factors (IL-3 and SCF). The resulting colonies are a mixture of mainly myeloid cells. The number and morphology of G418 resistant colonies can be determined after 2 weeks of selection. This strategy may be disadvantageous in that the appropriate stem cell may not divide and form G418 resistant colonies under the specific culture condition used. Moreover, it may be difficult to perform G418 selection on a population of heterogeneous cells, which vary in their sensitivity to G418. Therefore, another set of  $\Delta$ Ad.AAV<sup>fx</sup>-SNori infected CD34<sup>+</sup> cells can be cultured in methyl-cellulose (+IL-3, SCF) without G418 selection. After 2-3 weeks, single colonies can be picked from both (w/ and w/o G418) MC cultures, morphologically characterized, and analyzed for integrated vector using the modified protocol developed for integration studies in a small number of cells (see Figure 8). This strategy allows the assessment of whether hybrid vectors integrate into the genome of CD34<sup>+</sup> cells cultured in the presence of growth factors. This study gives us an idea about potential position effects affecting neo or  $\beta$ gal expression from integrated vector copies and about the structure of the integrated vector and the flanking chromosomal regions.

30 An alternative method to confirm vector integration:

Fluorescence *in situ* hybridization (FISH) analysis, can be performed in individual cells from MC colonies. CD34+ cells are cultured in MC in the presence of growth factors to induce cell division and subsequently treated with colchicine. Metaphase chromosome spreads are analyzed with biotin-ATP labeled probe specific for the  $\beta$ Gal or SNori gene and a dioxigenin-UTP labeled probe for the human X-chromosome as an internal control (provided by Christine Distech, University of Washington). Specific hybridization can be visualized with corresponding anti-biotin or anti-DIG antibodies labeled with different fluorochromes (e.g. FITC and Texas Red). Hybrid vector DNA may integrate as concatemers, which would facilitate detection by FISH. This technique allows one to localize the chromosomal integration sites of hybrid vectors.

Test transduction into quiescent bone marrow subpopulations:

Hybrid vectors described so far can be tested to see whether quiescent CD34+ cells can be stably transduced. To avoid significant cell proliferation, purified CD34+ cells are cultured in serum free IMDM supplemented with thrombopoietin (Tpo). Tpo can alone support the survival of stem cells without stimulating their active cell proliferation (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69). To analyze the proliferation status of CD34+ cells at the time point of infection with the hybrid vector  $\Delta$ Ad.AAV<sup>h</sup>BG, BrdU is added 2 hours before infection to the culture medium. One set of cells are maintained as suspension culture in IDAM containing Tpo only for two days. Another set of cells are grown in IDAM+Tpo supplemented with multiple cytokines. Forty eight hours after infection, CD34+ cells can be FACS sorted for beta Gal expression using FDG. FDG positive cells can be further analyzed for cellular DNA replication based on BrdU incorporation and for specific CD34+ subset markers. To do this, cytopins from FDG+ cells can be submitted to immunofluorescence with BrdU specific antibodies and with antibodies to specific cell surface markers (e.g. CD38, CD41). Alternatively, consecutive paraffin sections of the same cell can be analyzed for (a) transgene expression by X-Gal staining, (b) DNA synthesis based on BrdU incorporation, and (c) specific surface markers. This allows one to confirm that the culture conditions with Tpo alone prevent

significant genomic DNA replication and subsequent cell proliferation as well as to determine whether quiescent CD34+ cells can be infected based on beta Gal expression in cells where BrdU labeling is absent.

5 Test hybrid vectors integration into quiescent CD34+ cells:

Two sets of CD34+ cells are infected. The first set of  $\Delta$ Ad.AAV<sup>fx</sup>SNori infected cells are cultured for 5-7 days in the presence of cytokines; the other set is cultured without cytokines. To maintain CD34+ cell viability without cytokines during this period, the  
10 cells are cultured in the presence of Tpo or underlaid with a stromal cell line (AFT024) (Moore, K. et al., 1997, *Blood*, 89:4337-47), which can maintain HSC viable for 4 to 7 weeks. After this specific time period, both sets are submitted to clonogenic assays (in the presence of multiple cytokines) either in combination with G418 selection or without selection. Single colonies are analyzed morphologically and submitted to genomic DNA  
15 analysis (Figure 8) to determine the vector integration status. The ultimate proof for stem cell transduction is the in vivo survival/expansion assay. To do this, the CD34+ cells expressing beta Gal are used for transplantation experiments. If the number of FDG+ cells is not sufficient, total  $\Delta$ Ad.AAV<sup>fx</sup>BG infected cells as well as all  $\Delta$ Ad.AAV<sup>fx</sup>-SNori infected cells can be used directly without selection. Transplantation can be performed  
20 via tail vein injection into sublethally irradiated SCID NOD mice (Dao, M.A., et al., 1998, *Blood*, 4, 1243-1255; Matsunaga, T. et al., 1998, *Blood*, 92:452-61). At different time points after transplantation (4 to 8 weeks), mice can be sacrificed to obtain bone marrow cells which then can be cultured in suspension until various assays are performed for X-Gal and cell markers as described earlier. These cells also can be submitted to a  
25 secondary colony assay in MC or secondary transplantation into SCID NOD mice. Furthermore, MC colonies derived from these cells can be analyzed for the presence of integrated vector DNA by the method illustrated in Figure 8. The expression and integration data together allow conclusions about the repopulation efficiency and about potential position effects.

30

**B. Optimization of  $\Delta$ Ad.AAV<sup>fx</sup> vectors for  $\gamma$  r-globin expression in hematopoietic stem cells:**

One specific example of the invention is (a) to construct retargeted hybrid vector with the  $\gamma$ -globin as the transgene under the control of erythroid cell specific promoter, (b) to analyze the level and kinetics of  $\gamma$ -globin expression after transduction with hybrid vectors in *in vitro* and *in vivo* assays, (c) if required, to protect gene expression from position effects using  $\gamma$ -globin LCRs or insulators incorporated into hybrid vectors, and (d) to study whether  $\gamma$ -globin introns or heterologous introns can increase  $\gamma$ -globin expression.

Another central issue of the invention is to demonstrate that hybrid vectors can accommodate larger transgenes than rAAV and retroviruses. The insert size limitation of these vectors is 5kb. Transgene cassettes up to 8kb can be inserted into hybrid vectors as described. The maximal insert size may be about 14kb, if hybrid vectors are produced on the basis of E2a and/or E4 deleted rAd vectors in corresponding packaging cell lines. The maximal insert size in hybrid vectors is dictated by the packaging limit of first generation vectors (Ad.AAV) (<36kb) which are necessary intermediates for hybrid virus production at large scale. It is expected that stability and titer of Ad.AAV vectors with an 8kb globin gene cassette is comparable to the vector containing the 2.5-3.5kb cassette used in Ad.AAVBG, Ad.AAV1, and Ad.AAVSNori. The following example experiments address these issues.

**Production of  $\Delta$ Ad.AAV<sup>fx</sup> with large globin expression cassettes:**

In order to improve the condition of sickle cell disease, the expression level of the transferred  $\gamma$ -lobin gene must be at least 50% of that of each endogenous  $\beta$ gene. These levels of transgene expression can only be achieved by using optimal expression cassettes, including extended LCRs and intron containing gamma genes. (Forrester, W.C., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 1359-1363; Fraser, P., et al., 1998, *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in*

*Hematology*, 35, 105-111; Martin, D. et al., 1996, *Current Opinion in Genetics and Development*, 6:488-95), So far, most of the  $\gamma$ -globin expression cassettes are designed for retroviral and rAAV vectors, thus, less than 5kb and have to be devoid of internal splice sites or poly adenylation signals. With integrating vectors described herein, it is possible to go beyond this size limitation. This allows one to improve  $\gamma$ -globin expression in bone marrow cells in terms of an adequate expression level and long term persistence. For this purpose,  $\gamma$ -globin constructs developed by Li et al (Emery, D. W., et al. 1999 *Hum Gene Ther* 10:877-88; Li, Q., et al. 1999. *Blood* 93:2208-16) or by Ellis et al (Ellis, J., et al., 1996, *EMBO J.*, 15. 562-568; Ellis, J., et al., 1997, *Nucleic Acids Res.* 25, 1296-1302) is chosen.

(i) The first cassette contains a  $\gamma$ -globin expression unit used in retroviral vectors. This allows for a direct comparison between the two systems. This construct includes the beta promoter from -127 to the beta initiation codon, which is connected in frame with the gamma coding region. This beta promoter is combined with the 300bp HS40 derived from the human alpha globin locus, which acts as a strong enhancer for globin expression. The globin gene is the 1.1kb version with intron 1 and partially deleted intron 2. A second cassette is generated containing the HS40 beta promoter and gamma globin gene with the complete 3.3kb gamma globin gene.

20

(ii) The second construct contains the 6.5kb beta  $\mu$ LCR, which confer a dominant chromatin opening activity and an adequate level of gamma globin expression in transgenic mice. The LCR is linked to the short 1.1kb version of the gamma globin gene or the complete 3.3kb gamma gene.

25

(iii) Additional globin expression cassette can be generated which include insulators, MARs or SARs, as well as other elements that can improve transgene expression from integrated vectors or in transgenic animals, like introns derived from the HPRT or hGH genes (Chung, J.H., et al., 1997, *Proc. Natl. Acad. Sci. USA* 94, 575-580; Dunaway, M., et al, 1993, *Mol. Cell. Biol.*, 17, 182-189; Felsenfeld, G., et al., 1996,

30

*Proc. Natl. Acad. Sci. USA* 93, 93840-9388; Klehr, D., et al., 1991, *Biochemistry*, 30, 1264-1270).

Transduction studies with  $\Delta$ Ad.AAV $\gamma$ -globin vectors:

5

Transduction studies with globin-hybrid vectors are performed as described earlier (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13). Transduced CD34+ cells are submitted to differentiation in colony assays or analyzed *in vivo* expansion assays in SCID-NOD mice. MC-colonies or bone marrow cells from experimental mice are  
10 analyzed for globin expression. Gamma-globin expression is measured using fluorescent anti-gamma-globin antibodies. RNAase protection studies can be performed to specifically quantitate gamma globin mRNA in comparison with  $\gamma$ -globin RNA. For these studies around  $10^4$ - $10^5$  cells are needed per test.

15 Position effects:

In the absence of the LCR, globin genes are subjected to strong position effects when they are transferred into cultured CD34+ cells or erythroleukemic lines (Fraser, P., et al., 1998, *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in*  
20 *Hematology*, 35, 105-111). Another concern is that site-specific integration of  $\Delta$ Ad.AAV/rep vectors into AAVS1 may silence transgene expression. If silencing happens, it can be overcome by incorporating LCRs such as the 6.5kb  $\gamma$  globin  $\mu$ LCR (Ellis, J., et al., 1996, *EMBO J.*, 15, 562-568; Grosveld, F., et al., 1998, *Seminars in*  
25 *Hematology*, 35, 105-111) or insulators into  $\Delta$ Ad.AAV based expression units. Insulators are DNA elements that protect an integrated reporter gene from chromosomal position effects or that block enhancer activated transcription from a downward promoter. Insulator elements are known for *Drosophila melanogaster* genes (Gypsy, suppressor of Hairy wing, scs, scs', Fab-7). for the chicken beta-globin gene (HS4) and for the T cell receptor (BEAD1; 14, 21.25). Specifically, the *Drosophila* gypsy or the beta globin  
30 insulator can be inserted as two copies flanking the globin expression cassette into hybrid vectors. The position effects can be examined in transduced MC-colonies based on the



analysis of integrated vector DNA (see Fig. 29) and gamma-globin mRNA quantification. Analogous studies can be performed on transduced human bone marrow cells obtained after transplantation of infected CD34+ cells into SCID-NOD mice.

5 *Intron effects on gamma-globin expression:*

A number of reports reveal that the deletion of globin introns, particularly the second intron of the beta and gamma genes, decrease globin mRNA stability and thus the expression level (Antoniou, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9). RNA viruses  
10 such as onco-retro, lenti- and foami viruses are problematic as vehicles for intron-containing transgenes. Because  $\Delta$ Ad.AAV is a DNA virus, it should package globin introns and LCRs if necessary without the decreased titers and rearrangements observed with retroviral vectors.

## APPENDIX I

HUMAN AND ANIMAL ADENOVIRUSES AVAILABLE FROM  
AMERICAN TYPE CULTURE COLLECTION

- 1: Adenovirus Type 21 ATCC VR-1099 (NIAID V-221-002-014)
- 2: SA18 (Simian adenovirus 18) ATCC VR-943 Classification
- 3: SA17 (Simian adenovirus 17) ATCC VR-942 Classification
- 4: Adenovirus Type 47 ATCC VR-1309 Classification: Adenov
- 5: Adenovirus Type 44 ATCC VR-1306 Classification: Adenov
- 6: Avian adenovirus Type 4 ATCC VR-829 Classification: Ad
- 7: Avian adenovirus Type 5 ATCC VR-830 Classification: Ad
- 8: Avian adenovirus Type 7 ATCC VR-832 Classification: Ad
- 9: Avian adenovirus Type 8 ATCC VR-833 Classification: Ad
- 10: Avian adenovirus Type 9 ATCC VR-834 Classification: Ad
- 11: Avian adenovirus Type 10 ATCC VR-835 Classification: A
- 12: Avian adenovirus Type 2 ATCC VR-827 Classification: Ad
- 13: Adenovirus Type 45 ATCC VR-1307 Classification: Adenov
- 14: Adenovirus Type 38 ATCC VR-988 Permit: PHS permit requ
- 15: Adenovirus Type 46 ATCC VR-1308 Classification: Adenov
- 16: Simian adenovirus ATCC VR-541 Classification: Adenovir
- 17: SA7 (Simian adenovirus 16) ATCC VR-941 Classification:
- 18: Frog adenovirus (FAV-1) ATCC VR-896 Classification: Ad
- 19: Adenovirus type 48 (candidate) ATCC VR-1406 Classifica
- 20: Adenovirus Type 42 ATCC VR-1304 Classification: Adenov

- 21: Adenovirus type 49 (candidate) ATCC VR-1407 Classification:
- 22: Adenovirus Type 43 ATCC VR-1305 Classification: Adenovirus
- 23: Avian adenovirus Type 6 ATCC VR-831 Permit: USDA permit
- 24: Avian adenovirus Type 3 (Inclusion body hepatitis virus)
- 25: Bovine adenovirus Type 3 ATCC VR-639 Classification: A
- 26: Bovine adenovirus Type 6 ATCC VR-642 Permit: USDA permit
- 27: Canine adenovirus ATCC VR-800 Classification: Adenovirus
- 28: Bovine adenovirus Type 5 ATCC VR-641 Permit: USDA permit
- 29: Adenovirus Type 36 ATCC VR-913 Classification: Adenovirus
- 30: Ovine adenovirus type 5 ATCC VR-1343 Classification: A
- 31: Adenovirus Type 29 ATCC VR-272 Classification: Adenovirus
- 32: Swine adenovirus ATCC VR-359 Classification: Adenovirus
- 33: Bovine adenovirus Type 4 ATCC VR-640 Permit: USDA permit
- 34: Bovine adenovirus Type 8 ATCC VR-769 Permit: USDA permit
- 35: Bovine adenovirus Type 7 ATCC VR-768 Permit: USDA permit
- 36: Adeno-associated virus Type 2 (AAV-2H) ATCC VR-680 Classification:
- 37: Adenovirus Type 4 ATCC VR-4 Classification: Adenovirus
- 38: Adeno-associated virus Type 3 (AAV-3H) ATCC VR-681 Classification:
- 39: Peromyscus adenovirus ATCC VR-528 Classification: Adenovirus
- 40: Adenovirus Type 15 ATCC VR-661 Classification: Adenovirus
- 41: Adenovirus Type 20 ATCC VR-662 Classification: Adenovirus
- 42: Chimpanzee adenovirus ATCC VR-593 Classification: Adenovirus
- 43: Adenovirus Type 31 ATCC VR-357 Classification: Adenovirus
- 44: Adenovirus Type 25 ATCC VR-223 Classification: Adenovirus

- 45: Chimpanzee adenovirus ATCC VR-592 Classification: Aden
- 46: Chimpanzee adenovirus ATCC VR-591 Classification: Aden
- 47: Adenovirus Type 26 ATCC VR-224 Classification: Adenovi
- 48: Adenovirus Type 19 ATCC VR-254 Classification: Adenovi
- 49: Adenovirus Type 23 ATCC VR-258 Classification: Adenovi
- 50: Adenovirus Type 28 ATCC VR-226 Classification: Adenovi
- 51: Adenovirus Type 6 ATCC VR-6 Classification: Adenovirus
- 52: Adenovirus Type 2 Antiserum: ATCC VR-1079 AS/Rab (NIA
- 53: Adenovirus Type 6 ATCC VR-1083 (NIAID V-206-003-014)
- 54: Ovine adenovirus type 6 ATCC VR-1340 Classification: A
- 55: Adenovirus Type 3 ATCC VR-847 (Derived from NIAID V-20
- 56: Adenovirus Type 7 ATCC VR-7 Classification: Adenovirus
- 57: Adenovirus Type 39 ATCC VR-932 Classification: Adenovi
- 58: Adenovirus Type 3 ATCC VR-3 Classification: Adenovirus
- 59: Bovine adenovirus Type 1 ATCC VR-313 Classification: A
- 60: Adenovirus Type 14 ATCC VR-15 Classification: Adenovir
- 61: Adenovirus Type 1 ATCC VR-1078 (NIAID V-201-001-014)
- 62: Adenovirus Type 21 ATCC VR-256 Classification: Adenovi
- 63: Adenovirus Type 18 ATCC VR-1095 (NIAID V-218-003-014)
- 64: Baboon adenovirus ATCC VR-275 Classification: Adenovir
- 65: Adenovirus Type 10 ATCC VR-11 Classification: Adenovir
- 66: Adenovirus Type 33 ATCC VR-626 Classification: Adenovi

- 67: Adenovirus Type 34 ATCC VR-716 Classification: Adenovi
- 68: Adenovirus Type 15 ATCC VR-16 Classification: Adenovir
- 69: Adenovirus Type 22 ATCC VR-257 Classification: Adenovi
- 70: Adenovirus Type 24 ATCC VR-259 Classification: Adenovi
- 71: Adenovirus Type 17 ATCC VR-1094 (NIAID V-217-003-014)
- 72: Adenovirus Type 4 ATCC VR-1081 (NIAID V-204-003-014)
- 73: Adenovirus Type 16 ATCC VR-17 Classification: Adenovir
- 74: Adenovirus Type 17 ATCC VR-18 Classification: Adenovir
- 75: Adenovirus Type 16 ATCC VR-1093 (NIAID V-216-003-014)
- 76: Bovine adenovirus Type 2 ATCC VR-314 Classification: A
- 77: SV-30 ATCC VR-203 Classification: Adenovirus, Simian (
- 78: Adenovirus Type 32 ATCC VR-625 Classification: Adenovi
- 79: Adenovirus Type 20 ATCC VR-255 Classification: Adenovi
- 80: Adenovirus Type 13 ATCC VR-14 Classification: Adenovir
- 81: Adenovirus Type 14 ATCC VR-1091 (NIAID V-214-001-014)
- 82: Adenovirus Type 18 ATCC VR-19 Classification: Adenovir
- 83: SV-39 ATCC VR-353 Classification: Adenovirus, Simian (
- 84: Adenovirus Type 11 ATCC VR-849 (Derived from NIAID V-2
- 85: Duck adenovirus (Egg drop syndrome) ATCC VR-921 Permi "
- 86: Adenovirus Type 1 ATCC VR-1 Classification: Adenovirus
- 87: Chimpanzee adenovirus ATCC VR-594 Classification: Aden
- 88: Adenovirus Type 15 ATCC VR-1092 (NIAID V-215-003-014)

- 89: Adenovirus Type 13 ATCC VR-1090 (NIAID V-213-003-014)
- 90: Adenovirus Type 8 ATCC VR-1368 (Derived from NIAID V-20
- 91: SV-31 ATCC VR-204 Classification: Adenovirus, Simian (
- 92: Adenovirus Type 9 ATCC VR-1086 (NIAID V-209-003-014)
- 93: Mouse adenovirus ATCC VR-550 Classification: Adenoviru
- 94: Adenovirus Type 9 ATCC VR-10 Classification: Adenoviru
- 95: Adenovirus Type 41 ATCC VR-930 Classification: Adenovi
- 96: CI ATCC VR-20 Classification: Adenovirus, Simian (Mast
- 97: Adenovirus Type 40 ATCC VR-931 Classification: Adenovi
- 98: Adenovirus Type 37 ATCC VR-929 Classification: Adenovi
- 99: Marble spleen disease virus (Hemorrhagic enteritis virus
- 100: Adenovirus Type 35 ATCC VR-718 Classification: Adenovi
- 101: SV-32 (M3) ATCC VR-205 Classification: Adenovirus, Sim
- 102: Adenovirus Type 28 ATCC VR-1106 (NIAID V-228-003-014)
- 103: Adenovirus Type 10 ATCC VR-1087 (NIAID V-210-003-014)
- 104: Adenovirus Type 20 ATCC VR-1097 (NIAID V-220-003-014)
- 105: Adenovirus Type 21 ATCC VR-1098 (NIAID V-221-011-014)
- 106: Adenovirus Type 25 ATCC VR-1103 (NIAID V-225-003-014)
- 107: Adenovirus Type 26 ATCC VR-1104 (NIAID V-226-003-014)
- 108: Adenovirus Type 31 ATCC VR-1109 (NIAID V-231-001-014)
- 109: Adenovirus Type 19 ATCC VR-1096 (NIAID V-219-002-014)
- 110: SV-36 ATCC VR-208 Classification: Adenovirus, Simian (

- 111: SV-38 ATCC VR-355 Classification: Adenovirus, Simian (
- 112: SV-25 (M8) ATCC VR-201 Classification: Adenovirus, Sim
- 113: SV-15 (M4) ATCC VR-197 Classification: Adenovirus, Sim
- 114: Adenovirus Type 22 ATCC VR-1100 (NIAID V-222-003-014)
- 115: SV-23 (M2) ATCC VR-200 Classification: Adenovirus, Sim
- 116: Adenovirus Type 11 ATCC VR-12 Classification: Adenovir
- 117: Adenovirus Type 24 ATCC VR-1102 (NIAID V-224-003-014)
- 118: Avian adenovirus Type 1 (Chicken Embryo Lethal Orphan; C
- 119: SV-11 (M5) ATCC VR-196 Classification: Adenovirus, Sim
- 120: Adenovirus Type 5 ATCC VR-5 Classification: Adenovirus
- 121: Adenovirus Type 23 ATCC VR-1101 (NIAID V-223-003-014)
- 122: SV-27 (M9) ATCC VR-202 Classification: Adenovirus, Sim
- 123: Avian adenovirus Type 2 (GAL) ATCC VR-280 Classificati
- 124: SV-1 (M1) ATCC VR-195 Classification: Adenovirus, Simi
- 125: SV-17 (M6) ATCC VR-198 Classification: Adenovirus, Sim
- 126: Adenovirus Type 29 ATCC VR-1107 (NIAID V-229-003-014)
- 127: Adenovirus Type 2 ATCC VR-846 Classification: Adenovir
- 128: SV-34 ATCC VR-207 Classification: Adenovirus, Simian (
- 129: SV-20 (M7) ATCC VR-199 Classification: Adenovirus, Sim
- 130: SV-37 ATCC VR-209 Classification: Adenovirus, Simian (
- 131: SV-33 (M10) ATCC VR-206 Classification: Adenovirus, Si

- 132: Avian adeno-associated virus ATCC VR-865 Classification
- 133: Adeno-associated (satellite) virus Type 4 ATCC VR-646
- 134: Adenovirus Type 30 ATCC VR-273 Classification: Adenovi
- 135: Adeno-associated (satellite) virus Type 1 ATCC VR-645
- 136: Infectious canine hepatitis (Rubarth's disease, Fox ence
- 137: Adenovirus Type 27 ATCC VR-1105 (NIAID V-227-003-014)
- 138: Adenovirus Type 12 ATCC VR-863 (Derived from NIAID V-2
- 139: Adeno-associated virus Type 2 (molecularly cloned) ATCC
- 140: Adenovirus Type 7a ATCC VR-848 (Derived from NIAID V-2



What is claimed:

1. A first generation recombinant adenovirus vector a portion of which integrates into a host cell genome, comprising:
  - 5 a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
  - d) At least one adenoviral sequence which directs adenoviral replication; and
  - 10 e) A right adenoviral inverted terminal repeat sequence,wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.
2. The adenovirus vector of claim 1, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype.
3. The adenovirus vector of claim 1, wherein the sequence which directs adenoviral replication comprises a sequence on the anti-parallel strand which encodes an adenoviral fiber protein including a fiber tail, a fiber shaft, and a fiber knob, wherein the fiber knob includes a G-H loop region.
- 20 4. The adenovirus vector of claim 3, wherein the sequence on the anti-sense strand which encodes the fiber tail is from the same serotype as the adenoviral inverted repeat sequence.
5. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:
  - 30 a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;

- c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 5
6. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - b) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - 10 c) A transgene sequence 3' to the promoter sequence;
  - d) A polyadenylation sequence 3' to the transgene sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 15 7. The adenovirus vector of claim 5 or 6, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
8. The adenovirus vector of claim 5 or 6, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.
- 20
9. The adenovirus vector of claim 8, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.
- 25
10. The adenovirus vector of claim 8, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.
11. The adenovirus vector of claim 8, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
- 30

12. The transgene cassette of claim 5 or 6 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
- 5 13. The transgene cassette of claim 5 or 6, further comprising an insulator element.
14. The transgene cassette of claim 5 or 6, further comprising a bacterial origin of replication.
- 10 15. The adenoviral vector of claim 1, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
- 15 16. A first generation recombinant adenovirus vector which targets a host cell of interest and a portion of which integrates into the host cell genome so targeted, comprising two DNA strands, each strand being antiparallel to the other, the first strand comprising:
- 20 a) A left adenovirus inverted terminal repeat sequence;
- b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
- c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
- d) At least one adenoviral sequence which directs adenoviral replication; and
- 25 e) A right adenoviral inverted terminal repeat sequence,
- wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome, and wherein the second strand comprises a sequence which encodes an adenoviral fiber protein that permits targeting of the vector into the host cell of interest.

17. The adenovirus vector of claim 16, wherein the adenoviral protein includes a fiber tail, a fiber shaft, and a fiber knob, wherein the fiber knob includes a G-H loop region.
- 5 18. The adenovirus vector of claim 16, wherein the left and right adenovirus inverted terminal repeat sequences and the packaging sequence are from the same adenoviral serotype.
- 10 19. The adenovirus vector of claim 17, wherein the fiber tail is from the same serotype as the left and right adenoviral inverted repeat sequences.
20. The adenovirus vector of claim 17, wherein the fiber shaft is from a different serotype as the left and right adenoviral inverted repeat sequences.
- 15 21. The adenovirus vector of claim 20, wherein the fiber shaft is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
22. The adenovirus vector of claim 17, wherein the fiber shaft comprises a shortened length.
- 20 23. The adenovirus vector of claim 17, wherein the fiber knob is from a different serotype as the left and right adenoviral inverted repeat sequences.
24. The adenovirus vector of claim 23, wherein the fiber knob is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
- 25 25. The adenovirus vector of claim 17, wherein the fiber knob is a modified fiber knob protein comprising the G-H loop replaced with a heterologous peptide ligand sequence which binds to at least one surface protein on the host cell of interest.
- 30

26. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 5 b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - 10 e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
27. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 15 b) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the promoter sequence;
  - d) A polyadenylation sequence 3' to the transgene sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 20 27. The adenovirus vector of claim 26 or 27, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 25 28. The adenovirus vector of claim 26 or 27, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.
- 30 29. The adenovirus vector of claim 24, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin

30. The adenovirus vector of claim 24, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.
31. The adenovirus vector of claim 24, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
32. The transgene cassette of claim 26 or 27 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
33. The transgene cassette of claim 26 or 27, further comprising an insulator element.
34. The transgene cassette of claim 26 or 27, further comprising a bacterial origin of replication.
35. The adenoviral vector of claim 16, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
36. A recombinant gutless adenovirus vector a portion of which integrates into a host cell genome, comprising:
- a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence; and
  - d) A right adenoviral inverted terminal repeat sequence,
- wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.

37. The adenovirus vector of claim 36, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype. (*this is the base vector*)
- 5 38. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - 10 c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 15 39. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - b) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the promoter sequence;
  - 20 d) A polyadenylation sequence 3' to the transgene sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
40. The adenovirus vector of claim 38 or 39, wherein the left and the right inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 25 41. The adenovirus vector of claim 38 or 39, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.
- 30

42. The adenovirus vector of claim 41, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin
43. The adenovirus vector of claim 41, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentomycin.
44. The adenovirus vector of claim 41, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
45. The adenovirus vector of claim 38 or 39 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
46. The adenovirus vector of claim 38 or 39, further comprising an insulator element.
47. The adenovirus vector of claim 38 or 39, further comprising a bacterial origin of replication.
48. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 5' portion of a gene of interest.
49. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 3' portion of a gene of interest.
50. A method of producing a resolved gutless adenovirus vector in a suitable cell, said method comprising introducing a first and a second adenovirus vectors of claim 1 or 16 into the cell under suitable conditions so that the recombinant adenovirus vectors undergo homologous recombination thereby producing a resolved gutless adenovirus vector.



51. A resolved gutless adenovirus vector produced by the method of claim 50.
52. The method of claim 50, wherein the first adenovirus vector comprises a transgene cassette having a 5' portion of a gene of interest, and wherein the  
5 second adenovirus vector comprises a transgene cassette having a 3' portion of the gene of interest, and wherein a part of the 5' portion overlaps with a part of the 3' portion so that homologous recombination occurs.
53. A method of producing a resolved gutless recombinant Ad vector by homologous  
10 recombination in a suitable cell, said method comprising contacting two parental recombinant Ad vectors, each comprising a transgene cassette containing a portion of a selected transgene with a region of overlapping homology, so that the first and second parental recombinant Ad vectors undergo homologous recombination at the region of overlapping homology, resulting in a resolved  
15 recombinant gutless Ad vector having both portions of the selected transgene, and wherein the selected transgene is within a transgene cassette flanked by a pair of ITRs.
54. A resolved gutless adenovirus vector produced by the method of claim 53.
- 20 55. An adenovirus library comprising a plurality of adenovirus vectors expressing fiber proteins which are displayed and modified by random peptide insertions.
56. The library of claim 55, wherein said fiber protein so displayed comprises a  
25 random peptide substituted in the G-H loop of the fiber protein knob domain.
57. A screening method for targeting adenovirus vectors for gene therapy comprising  
30 contacting the adenovirus library of claim 55 with a plurality of cells so that the cells are transduced with the adenovirus vectors of the adenovirus library transduction occurs and detecting the cells so transduced.

58. The adenovirus vector of claim 5, 6, 26, 27, 38, or 39 further comprising a nucleotide sequence encoding a rep78 protein.

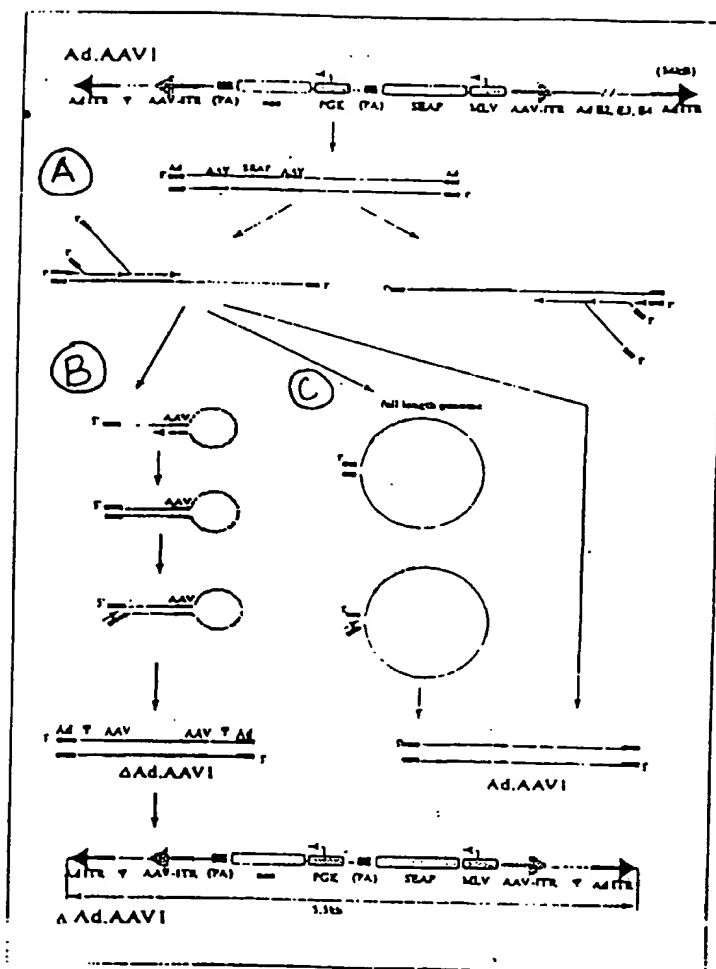


FIG. 1

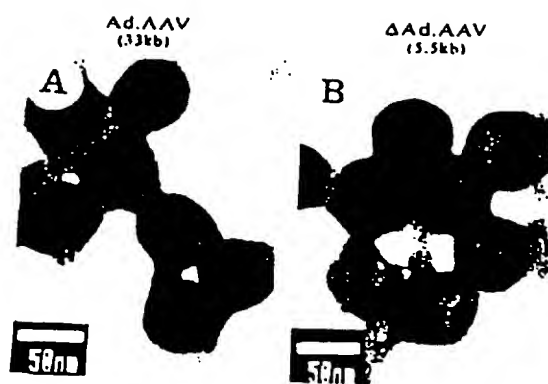


FIG. 2



FIG. 3

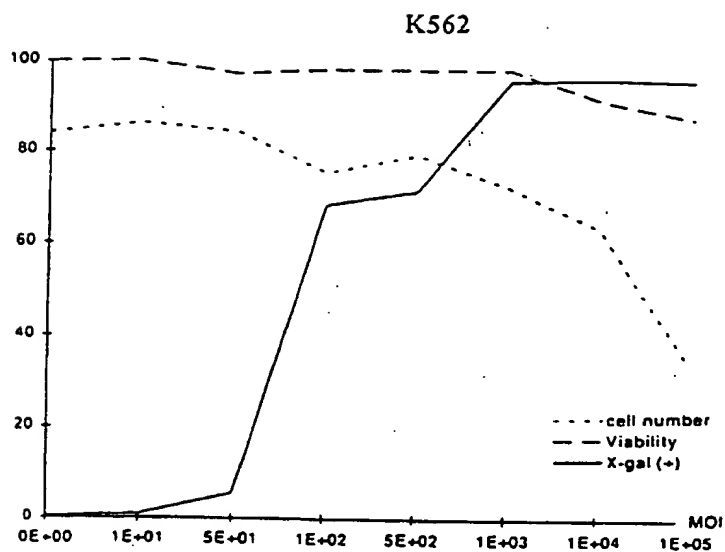


FIG. 4A

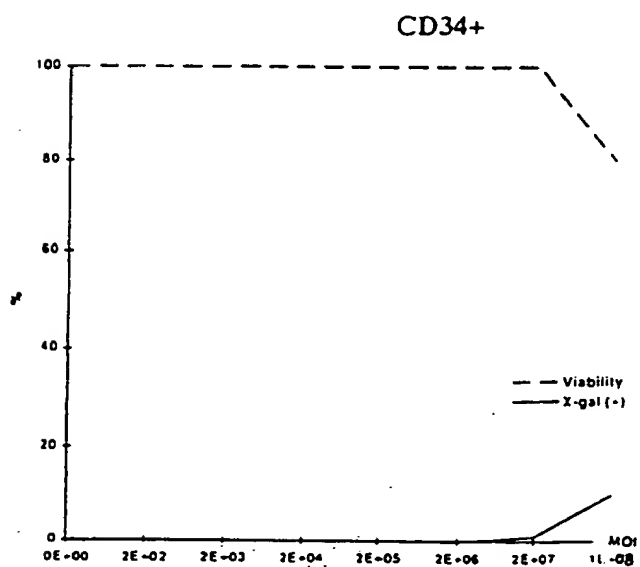


FIG. 4B

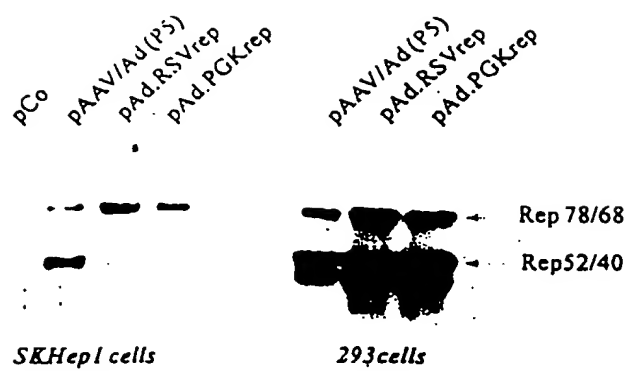


FIG. 5

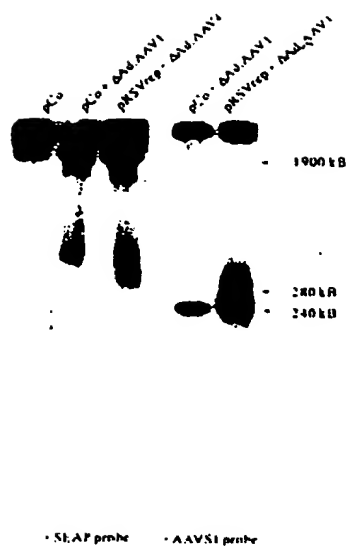


FIG. 6



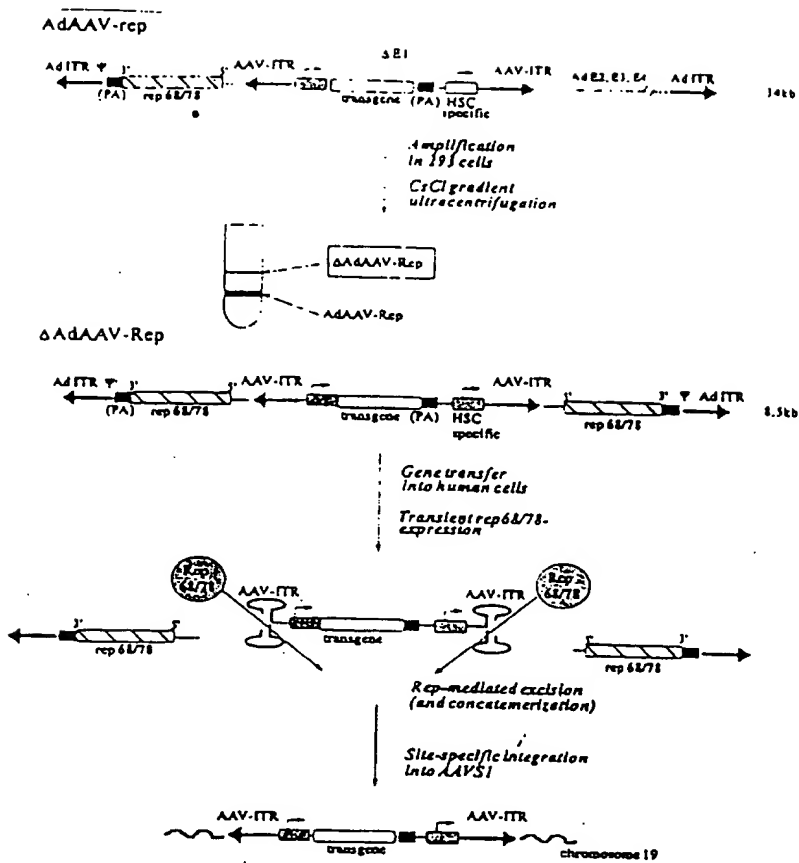


FIG. 7



FIG. 8A

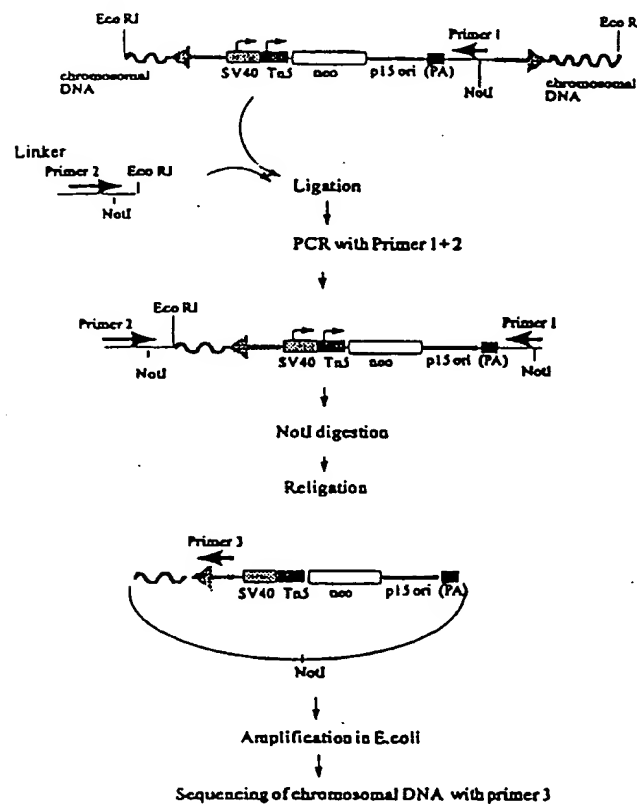
Integrated  $\Delta$ Ad.AAV-SNori

FIG. 8B

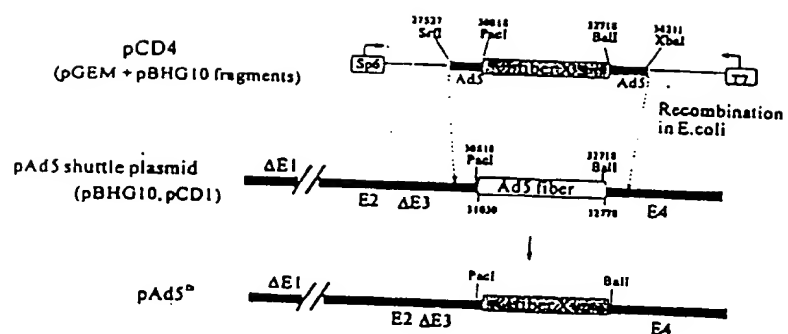


FIG. 9

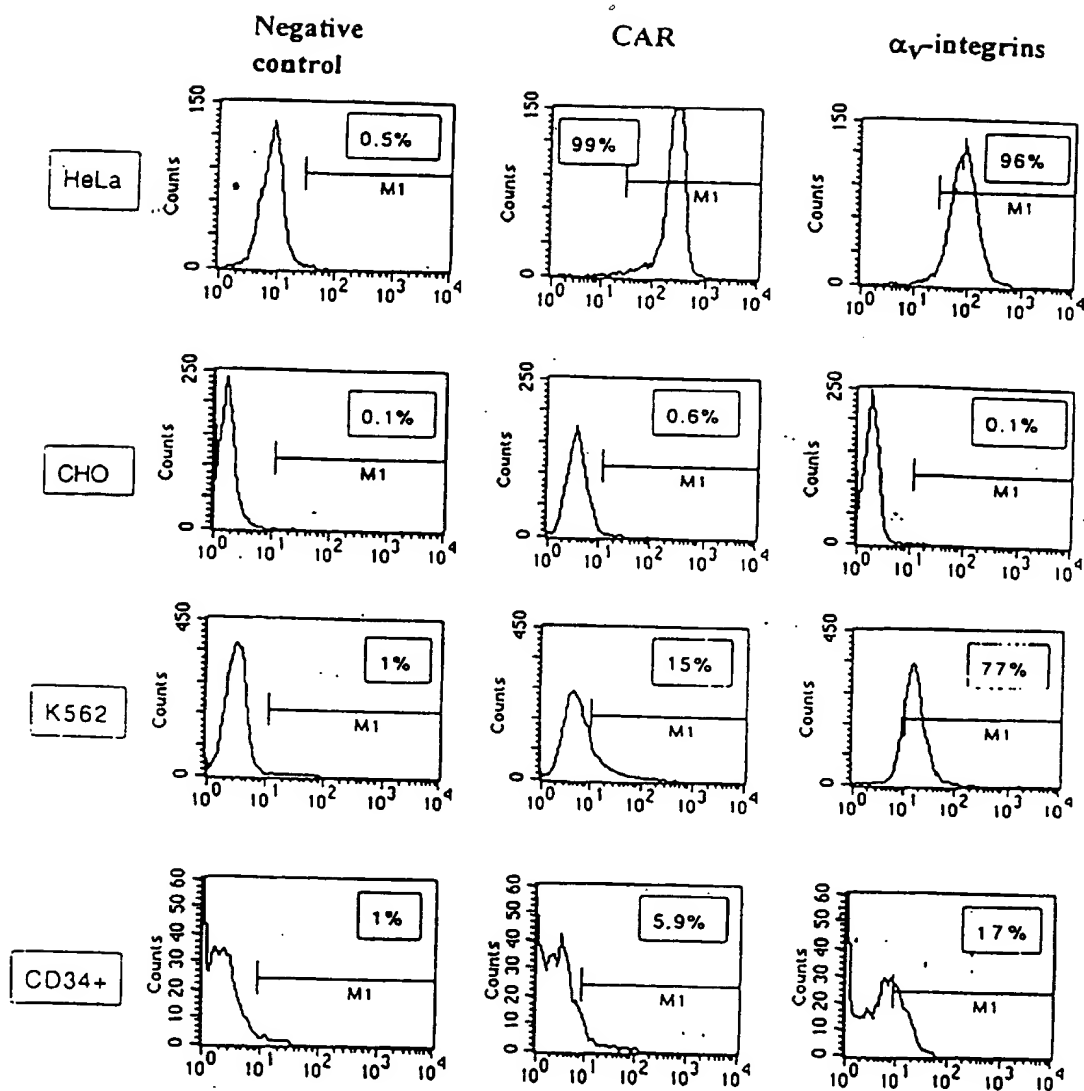


FIG. 10

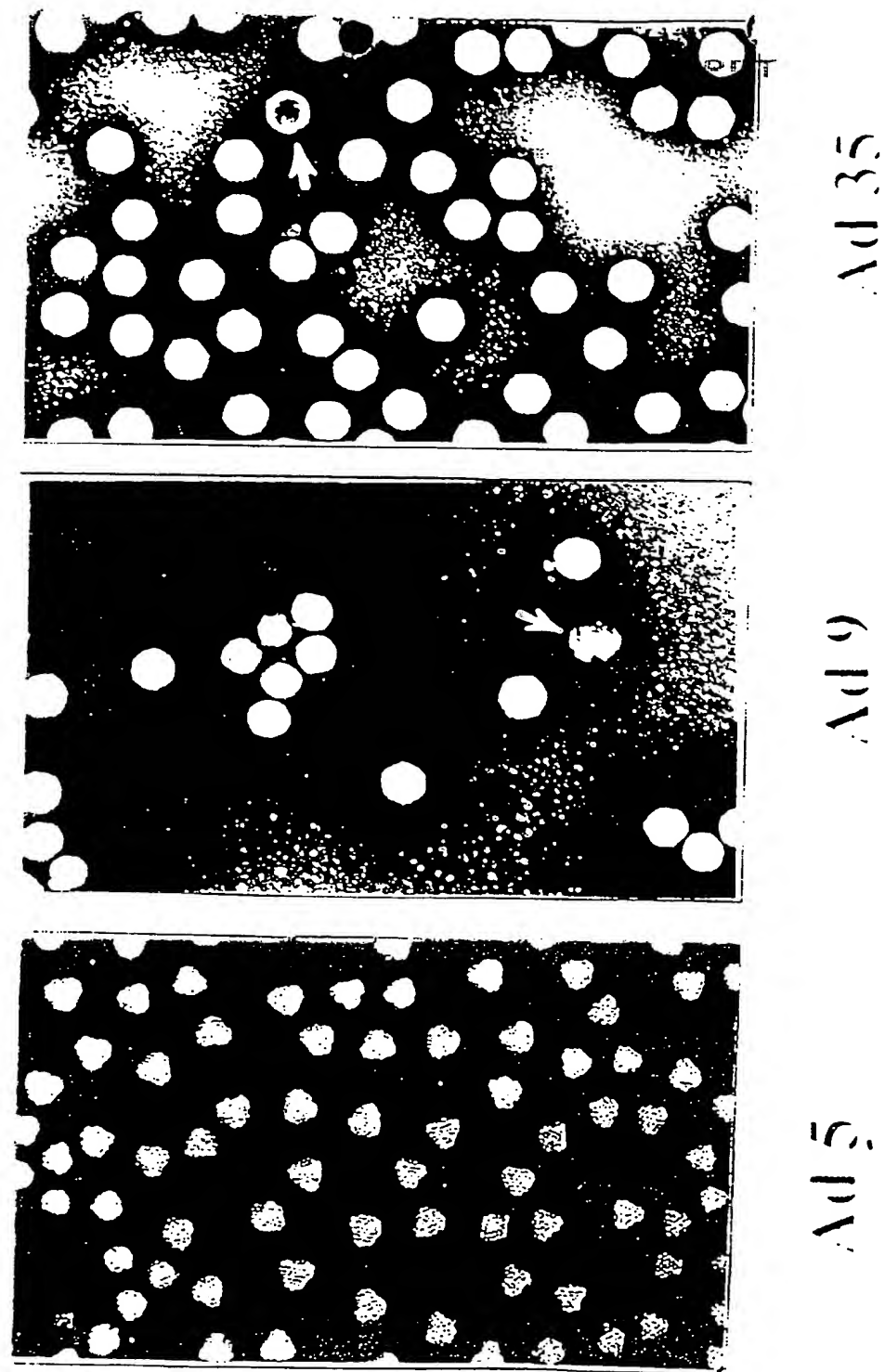


Fig. 11

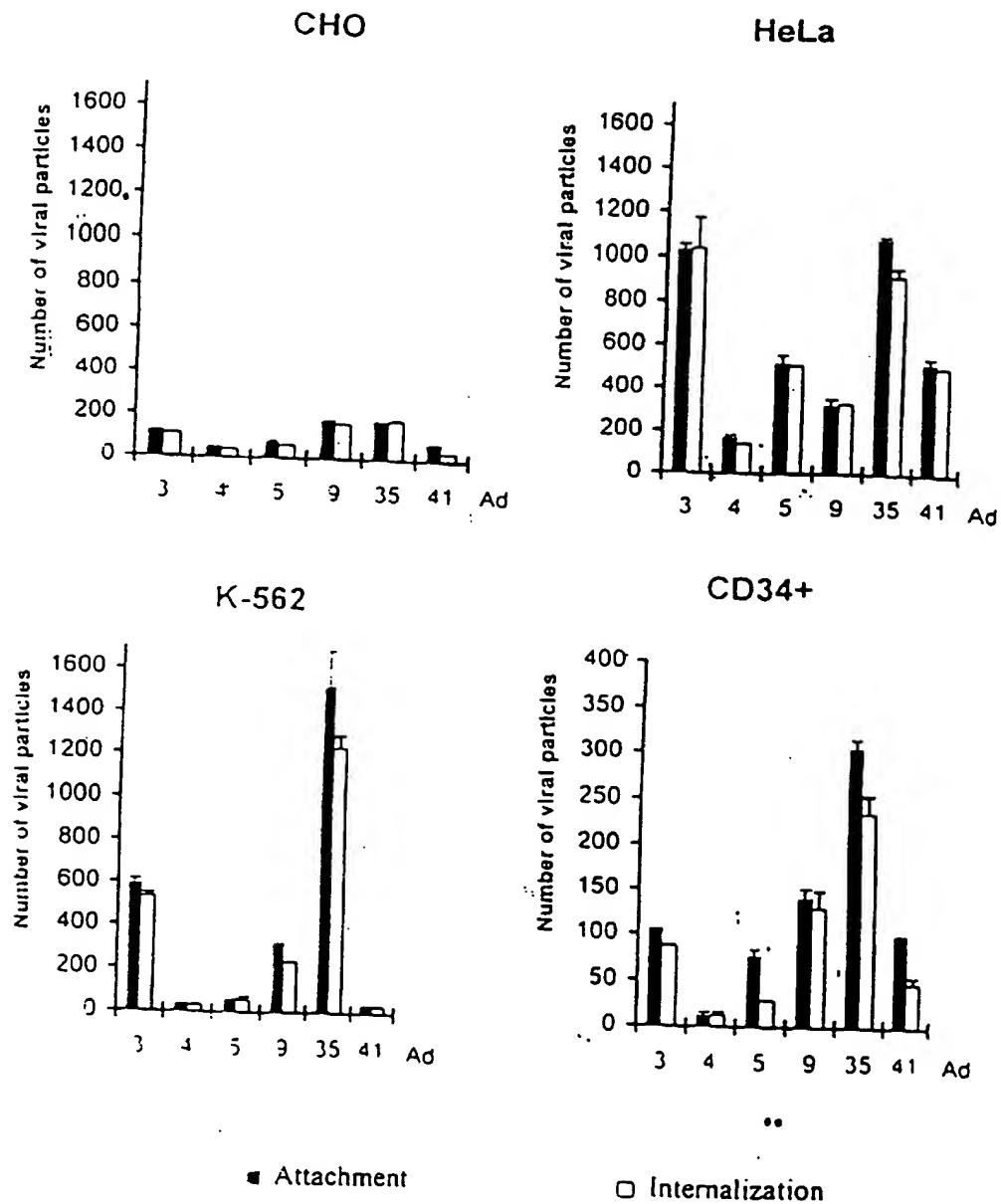


FIG. 12

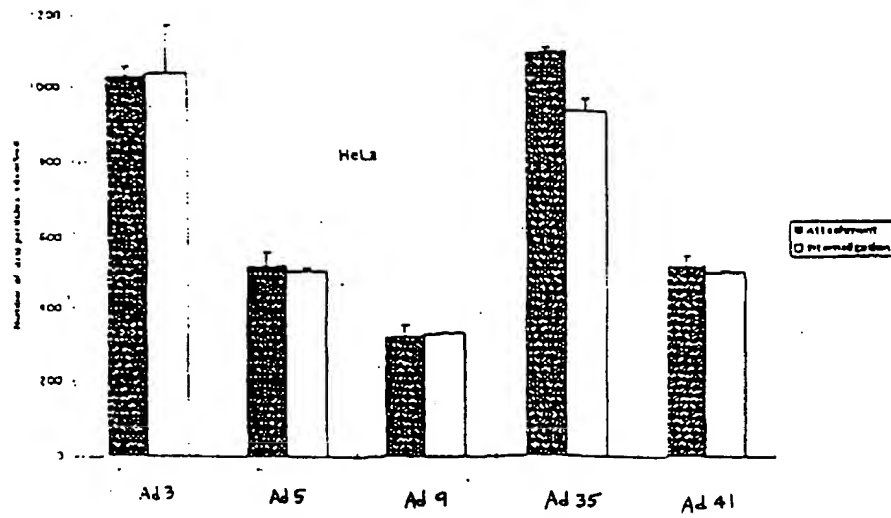


FIG. 13 A

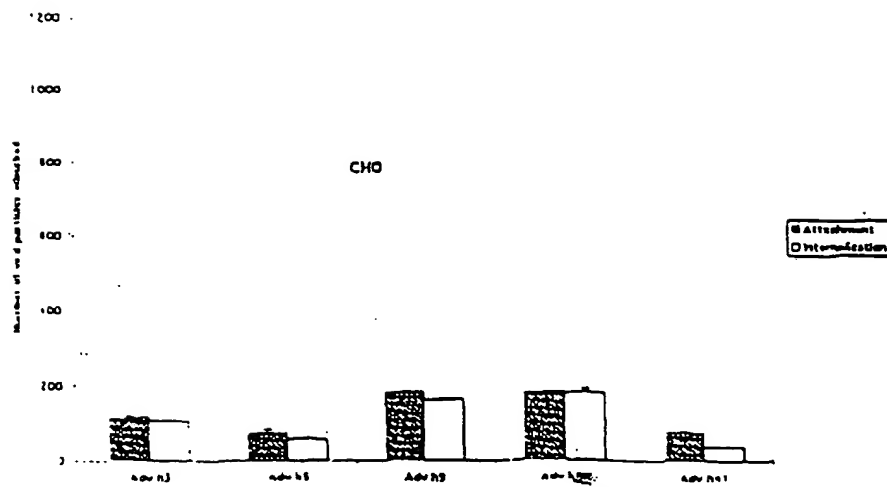


FIG. 13 B

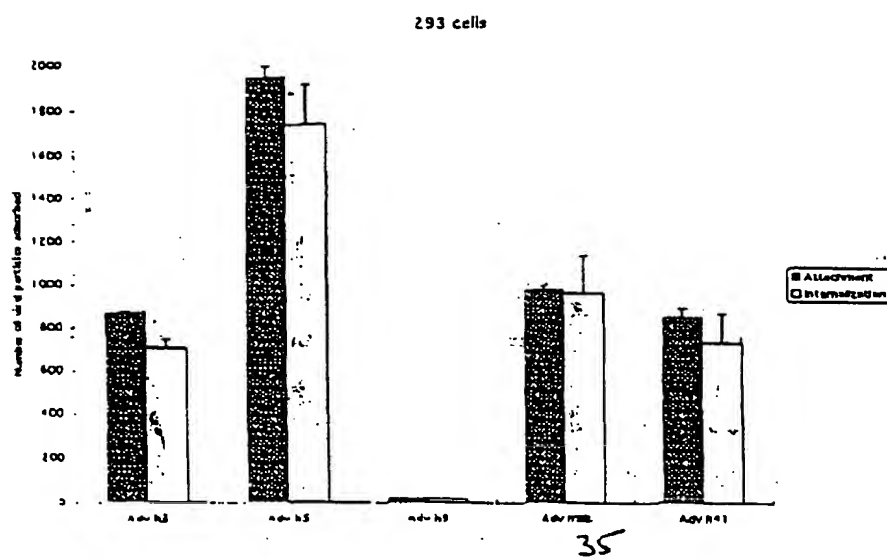


FIG. 13C



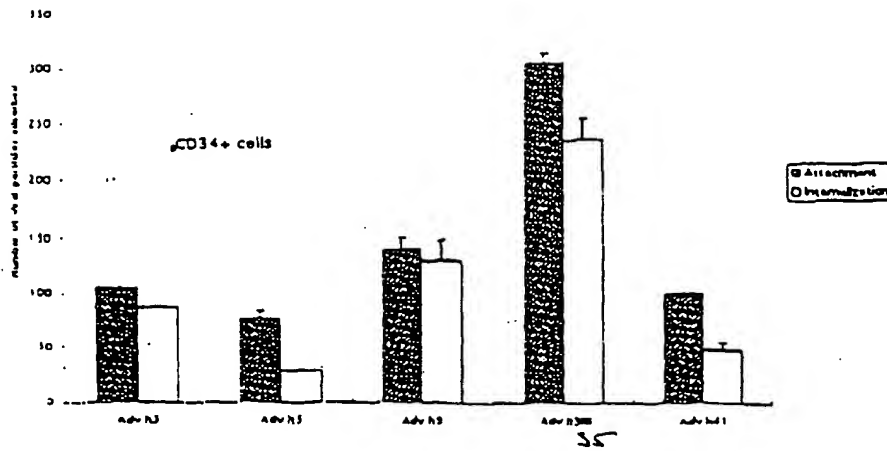


FIG 14A

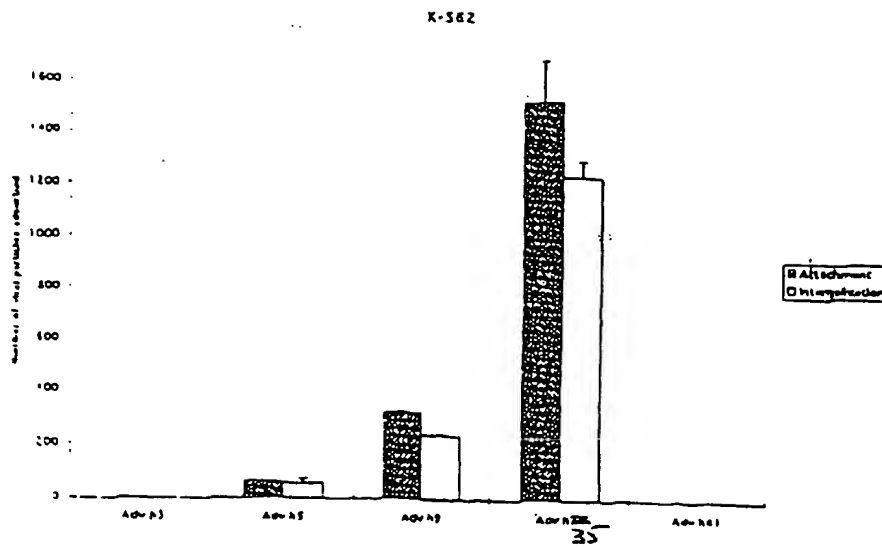


FIG 14B

A. K562 cells

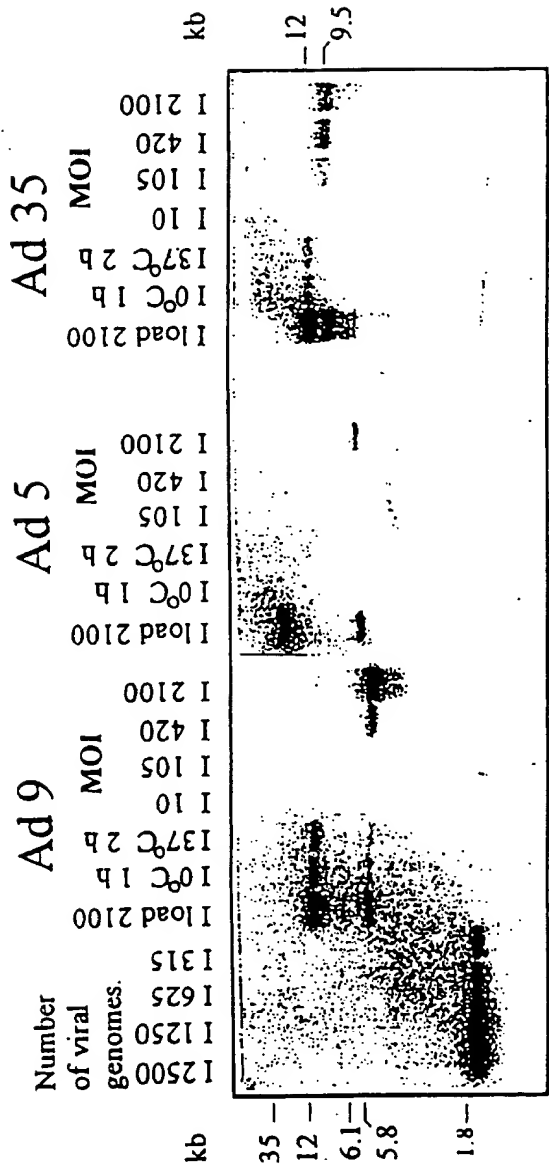
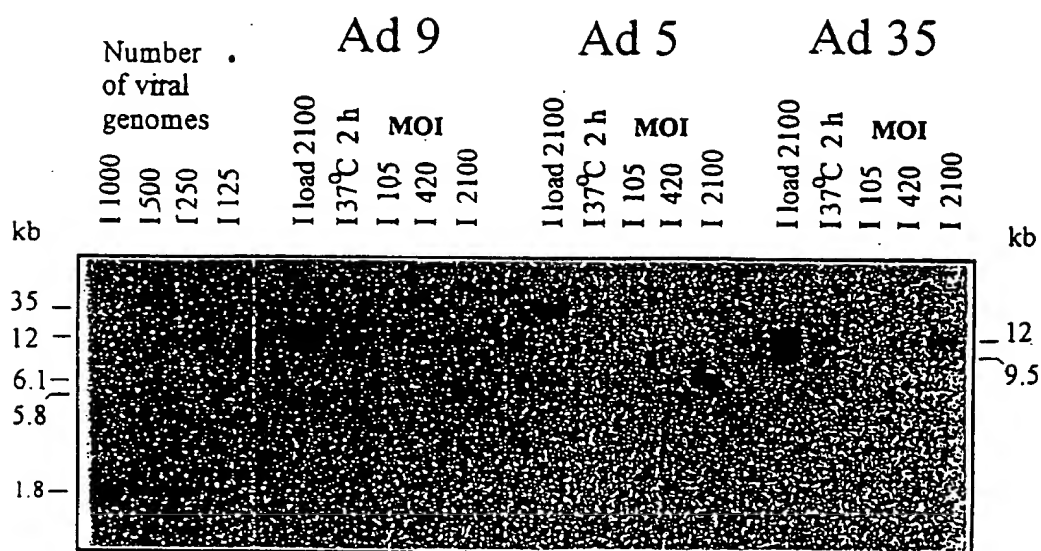


FIG. 15

## B. CD 34+ cells



## C.

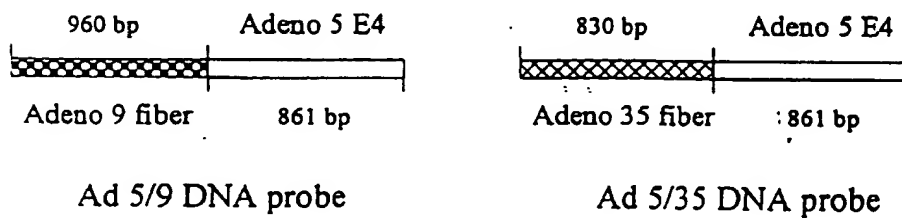


FIG. 15 continued

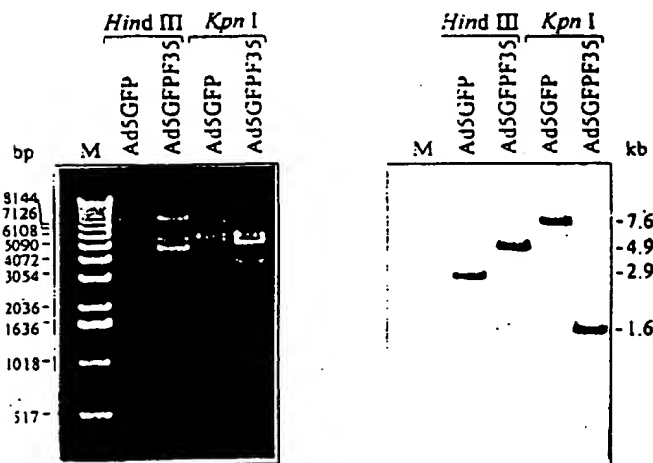
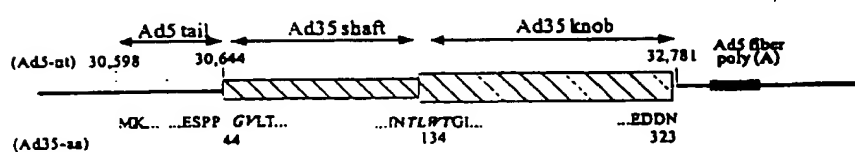
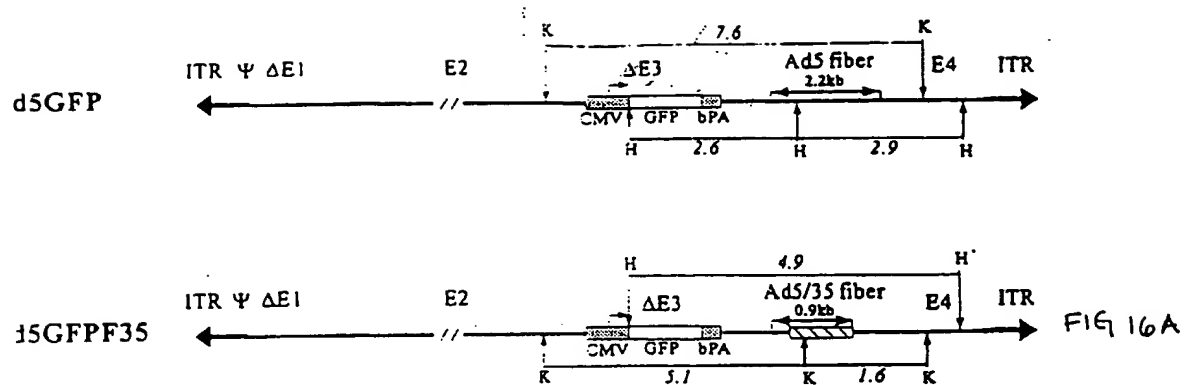


FIG. 16B

FIG. 16

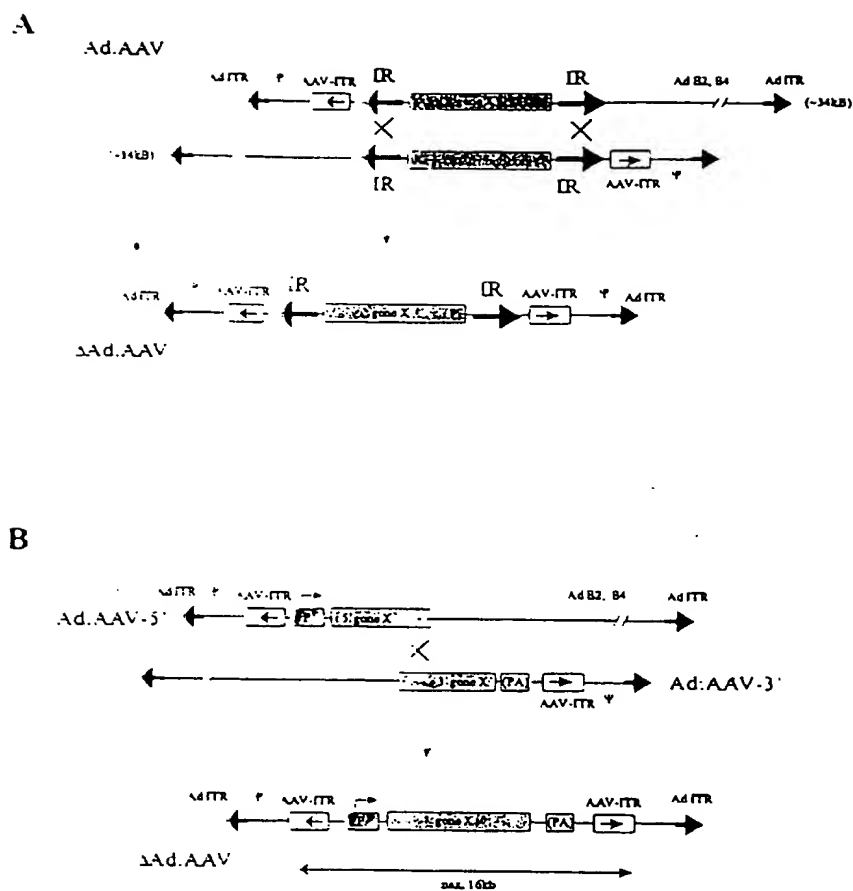


FIG. 17

## Ad5.GFP/F35

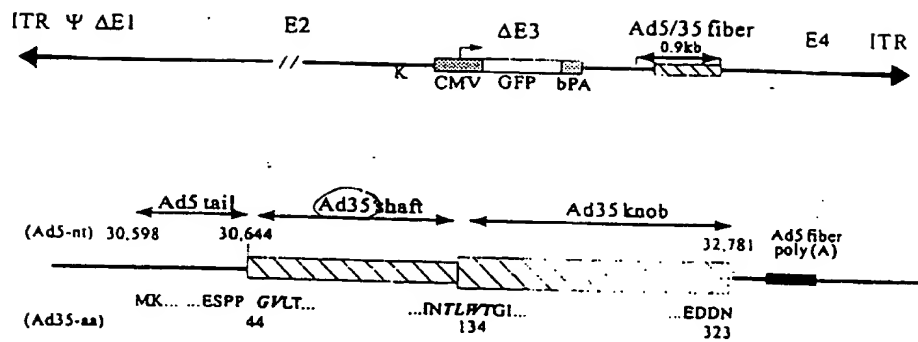


FIG. 18

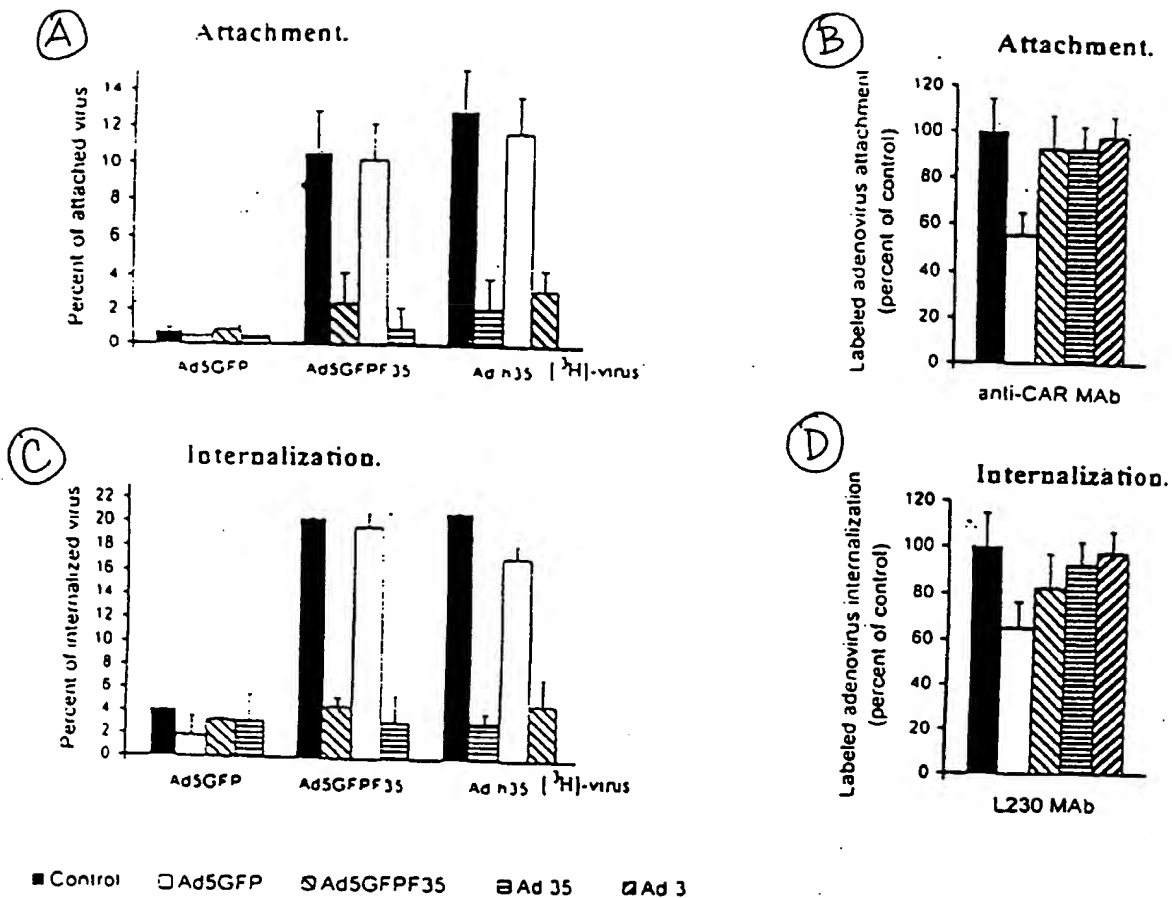


FIG. 19

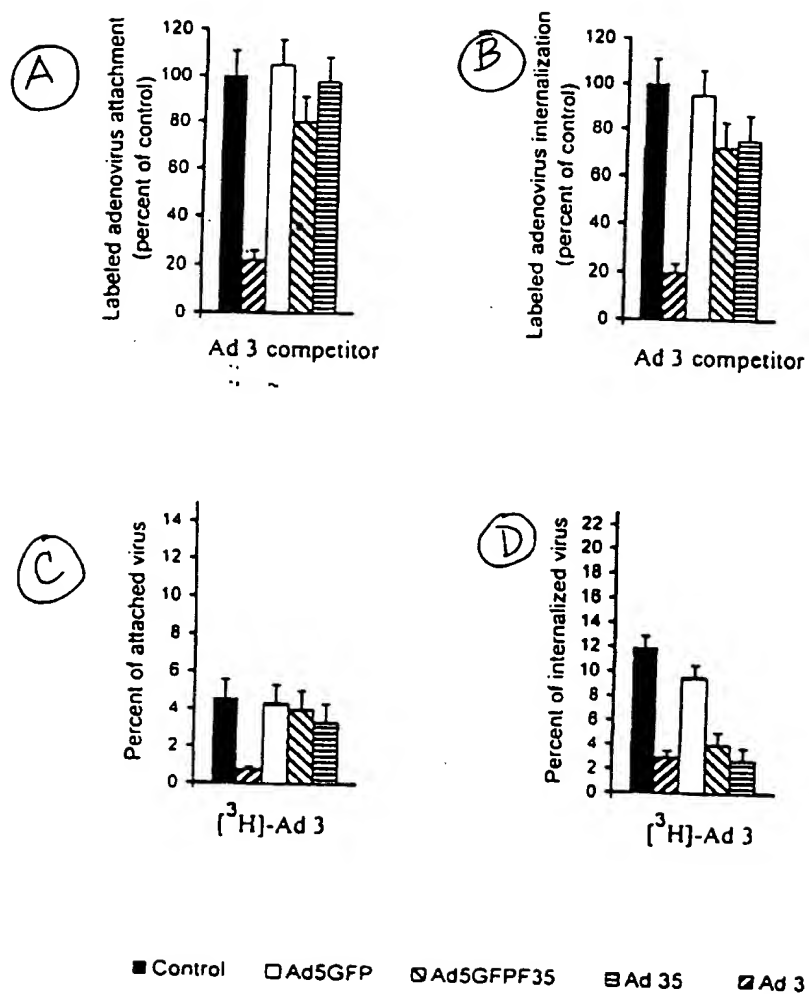


FIG. 20



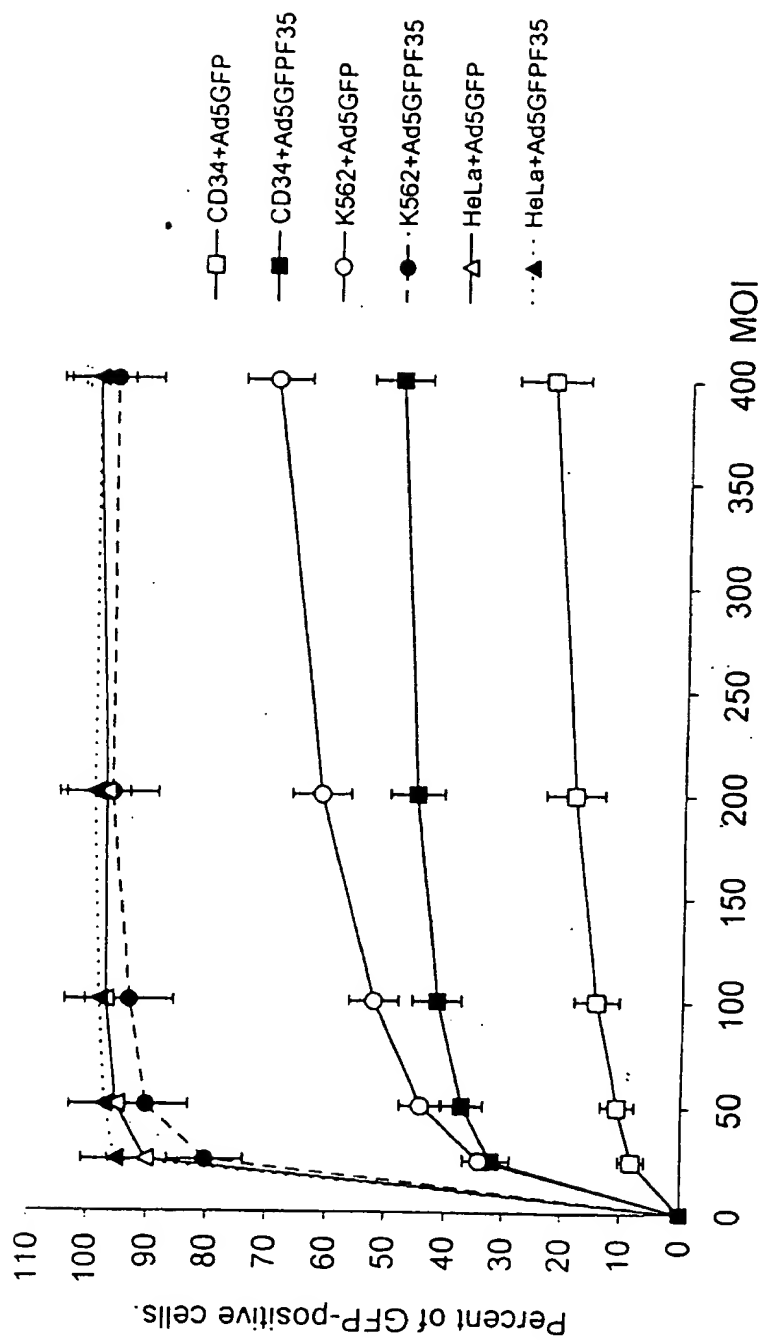


FIG. 21

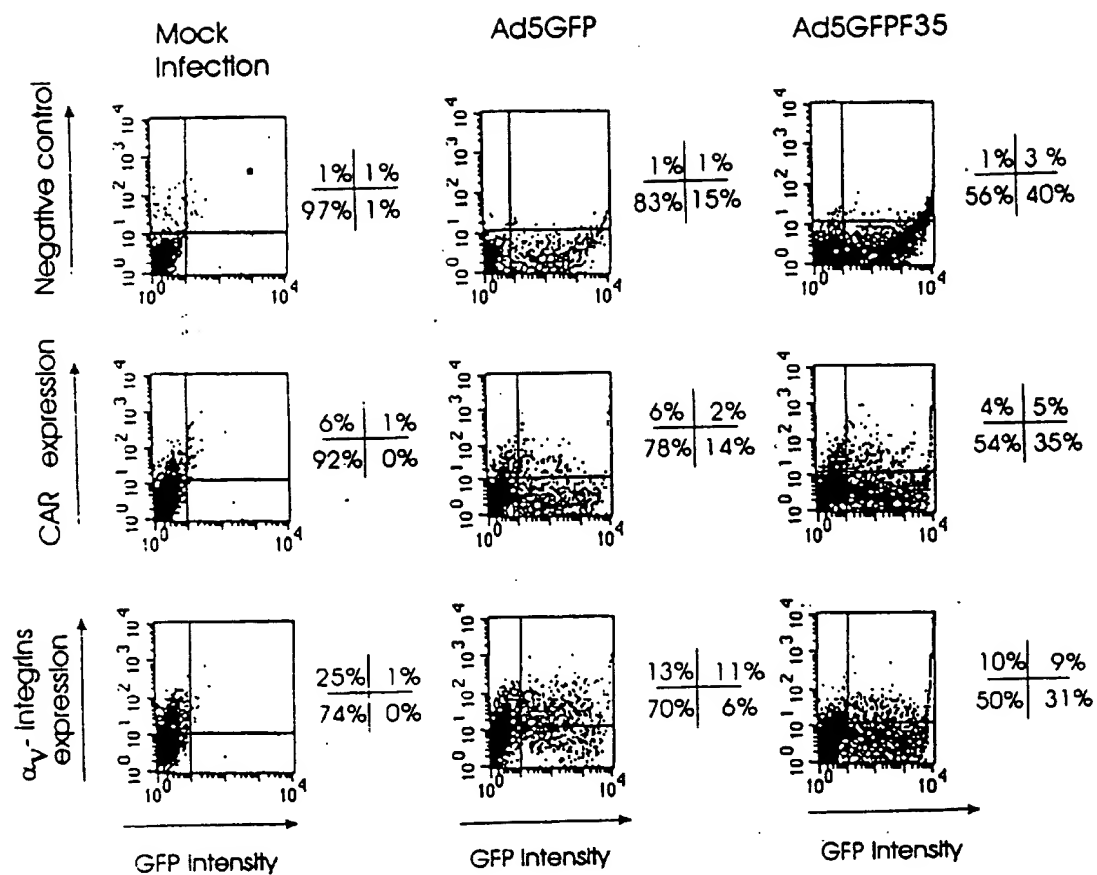


FIG. 22

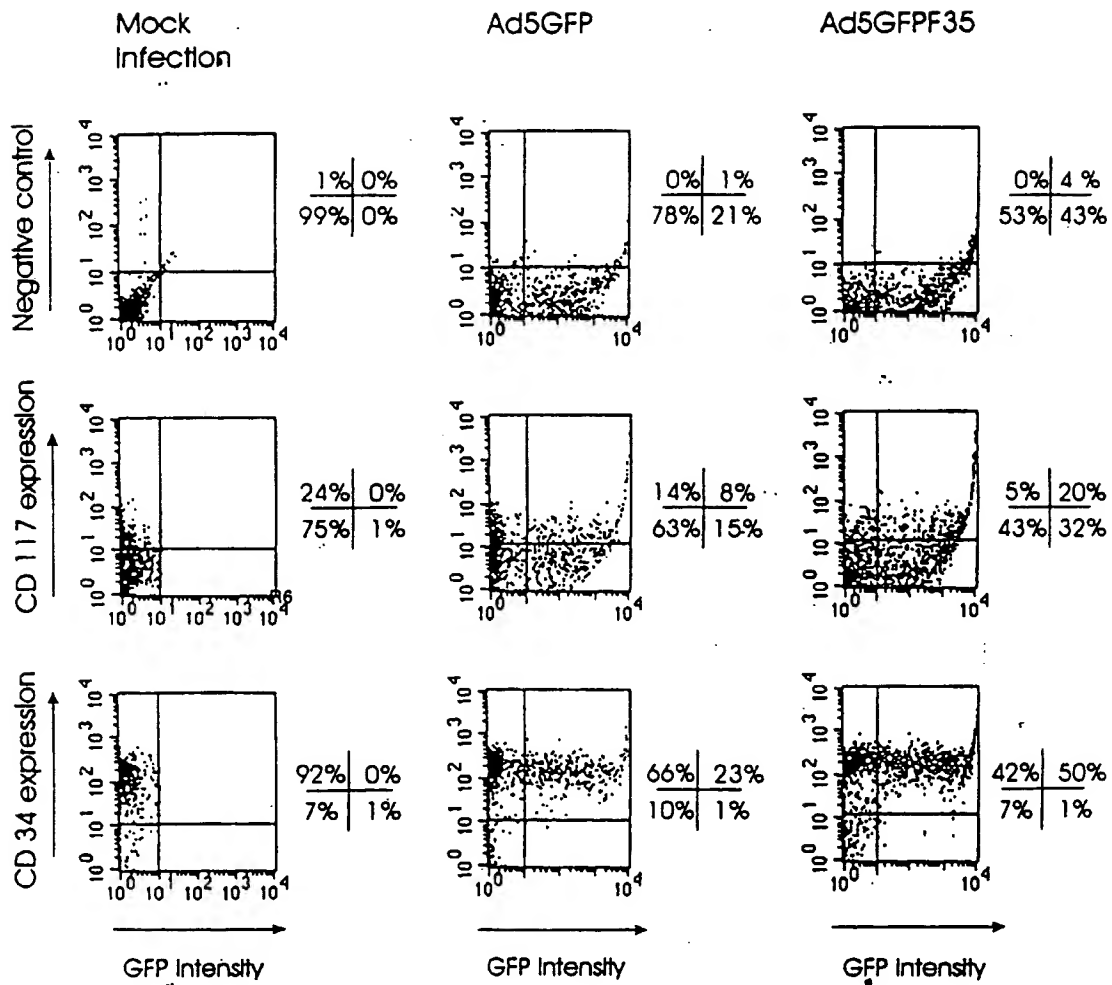


FIG. 23 A

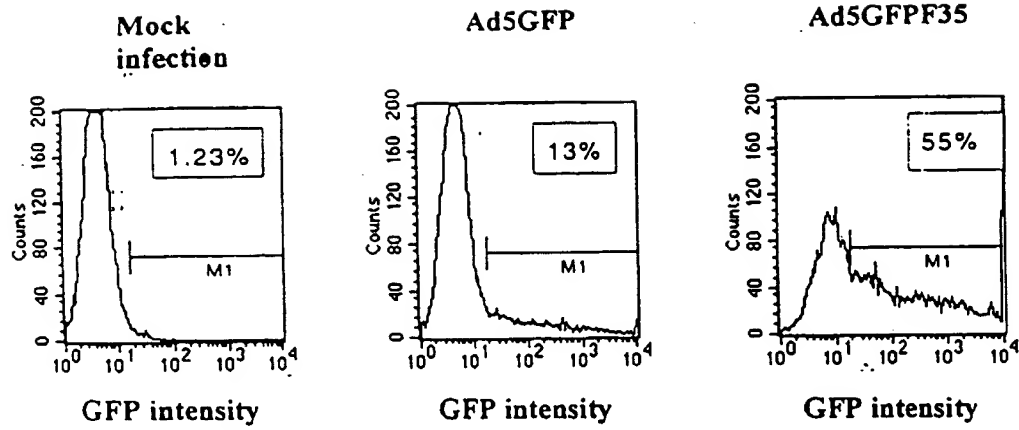


FIG 23 B

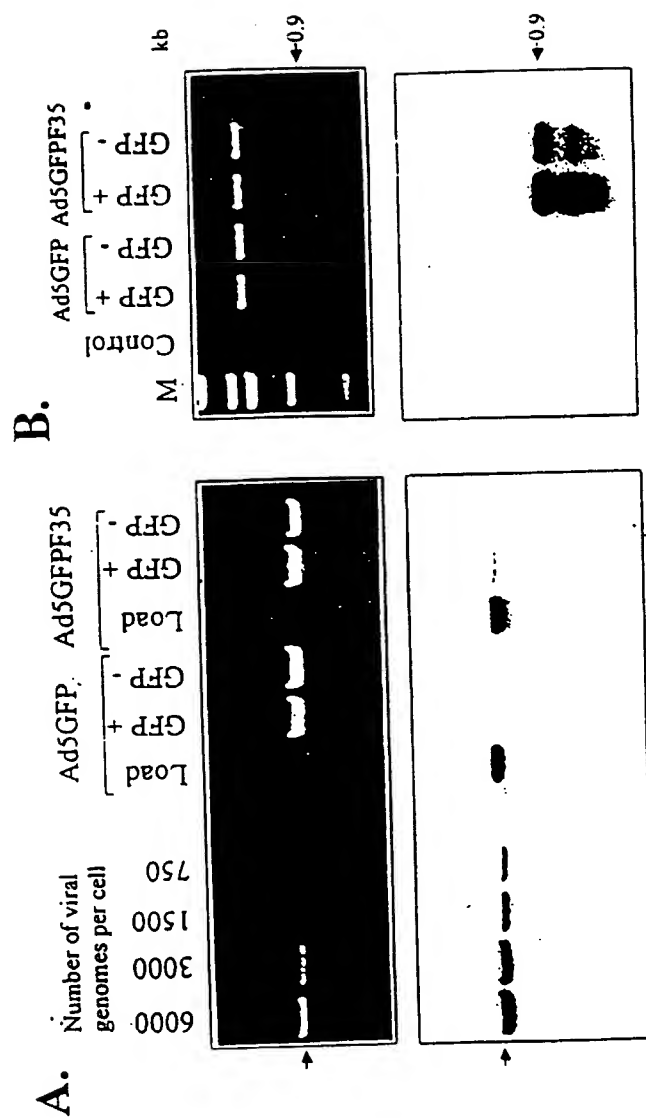


FIG. 24

Structure and features of chimeric adenovirus vectors

| Adeno vector                        | Ad 5L | Ad 5S | Ad 5/9L | Ad 5/9S | Ad 5/35L | Ad 5/35S |
|-------------------------------------|-------|-------|---------|---------|----------|----------|
| Binding to CAR                      | Yes   | Yes   | Yes     | Yes     | No       | No       |
| Genome/PPU ratio                    | 13    | 143   | 29      | 582     | 17       | 14       |
| Attachment to/ internalization into |       |       |         |         |          |          |
| 293                                 | +++   | ++    | ++      | +/-     | +++      | +++      |
| K562                                | -/+   | -/+   | -/+     | ---     | +++      | +++      |
| Y79                                 | ++    | +     | +       | +/-     | +        | +        |
| Gene transfer efficiency            |       |       |         |         |          |          |
| 293                                 | +++   | ++    | ++      | +/-     | +++      | +++      |
| K562                                | -/+   | -/+   | -/+     | ---     | +++      | +++      |
| Y79                                 | ++    | +     | +       | +/-     | +        | +        |

FIG. 25

# Top view of Adenovirus type 5 fiber knob domain

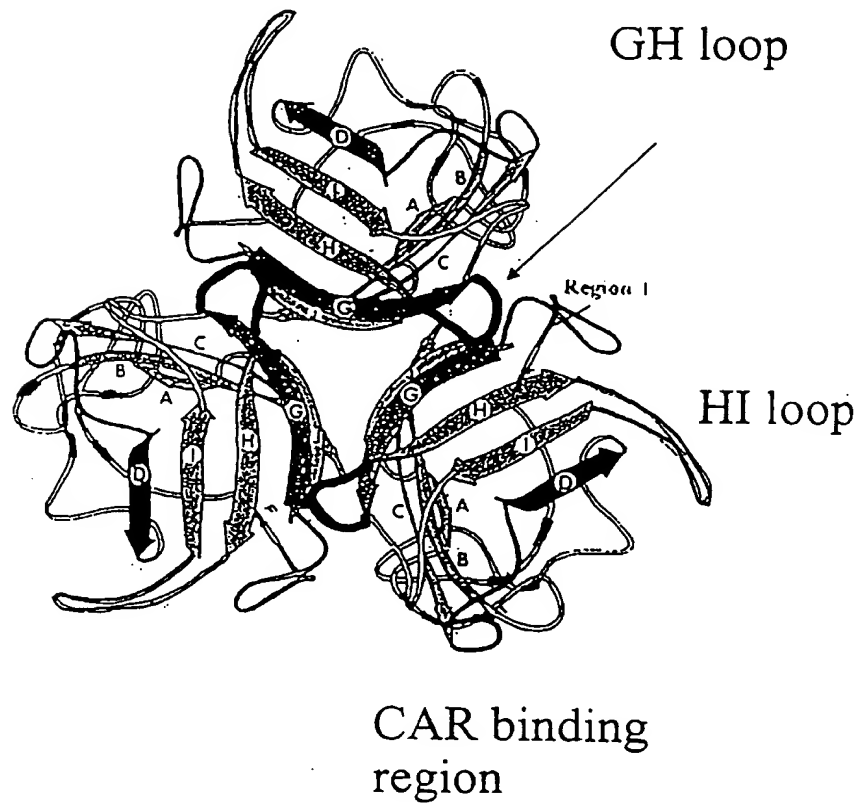


FIG. 26

## Amino acids composition of wt and modified Adeno 5 GH loop

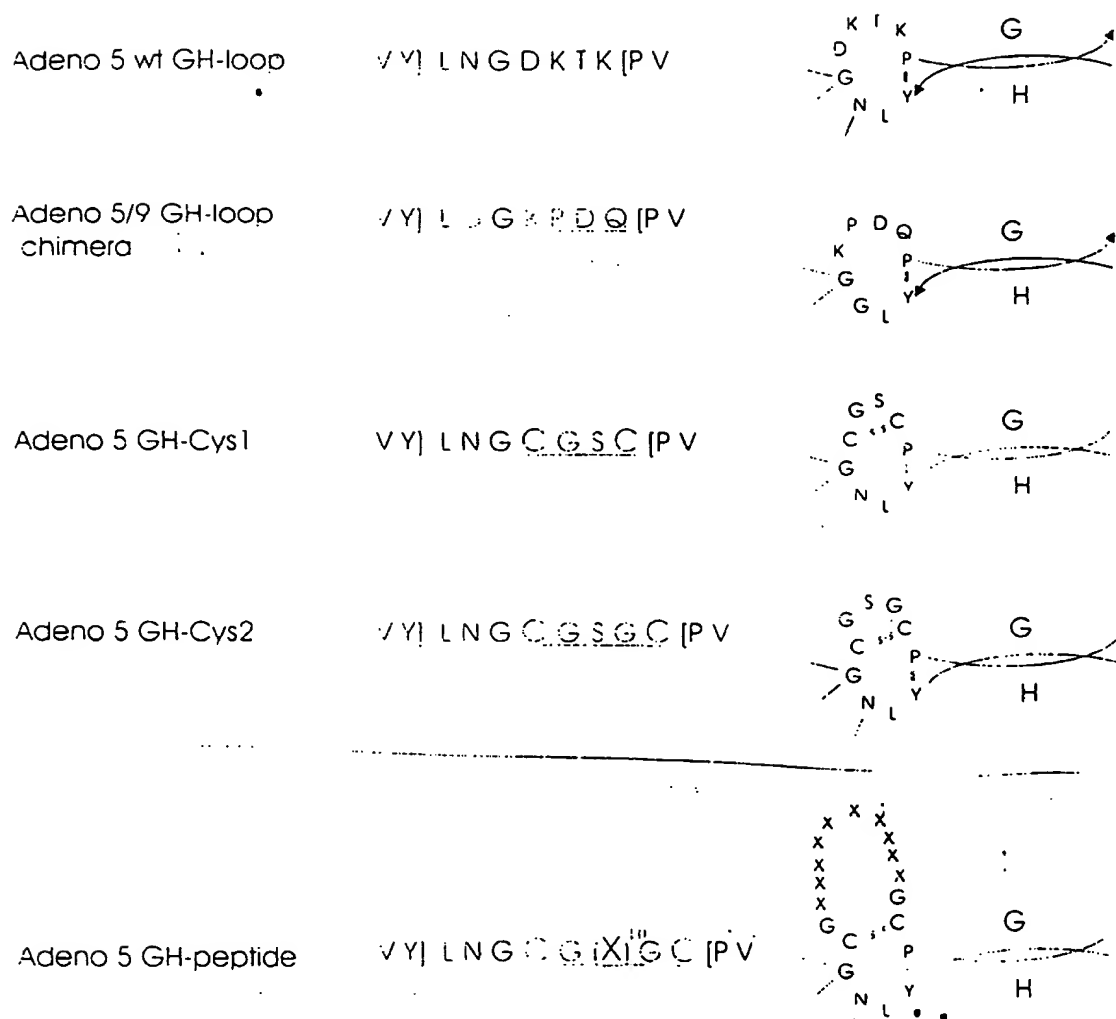


FIG. 27



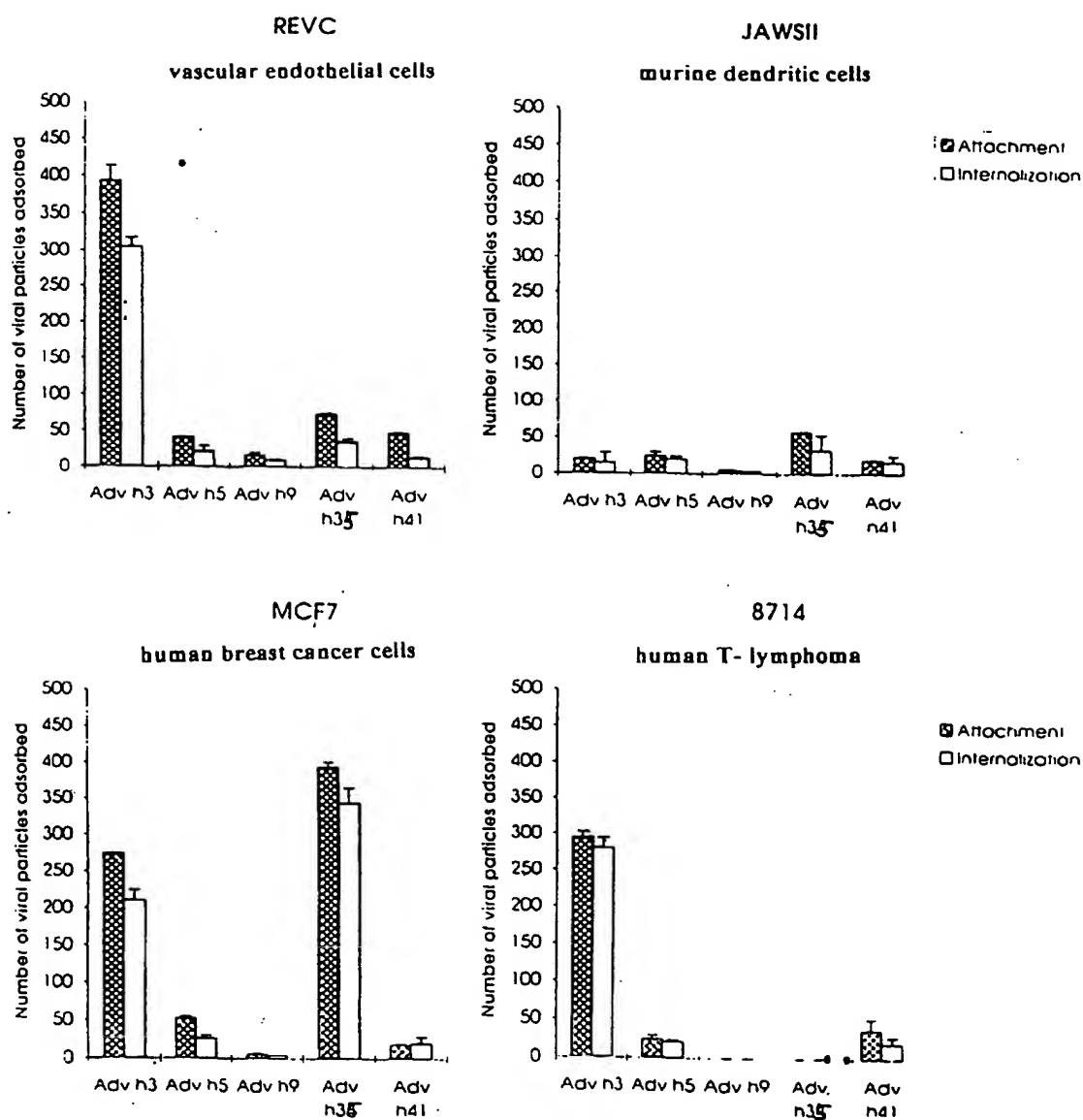


FIG 28

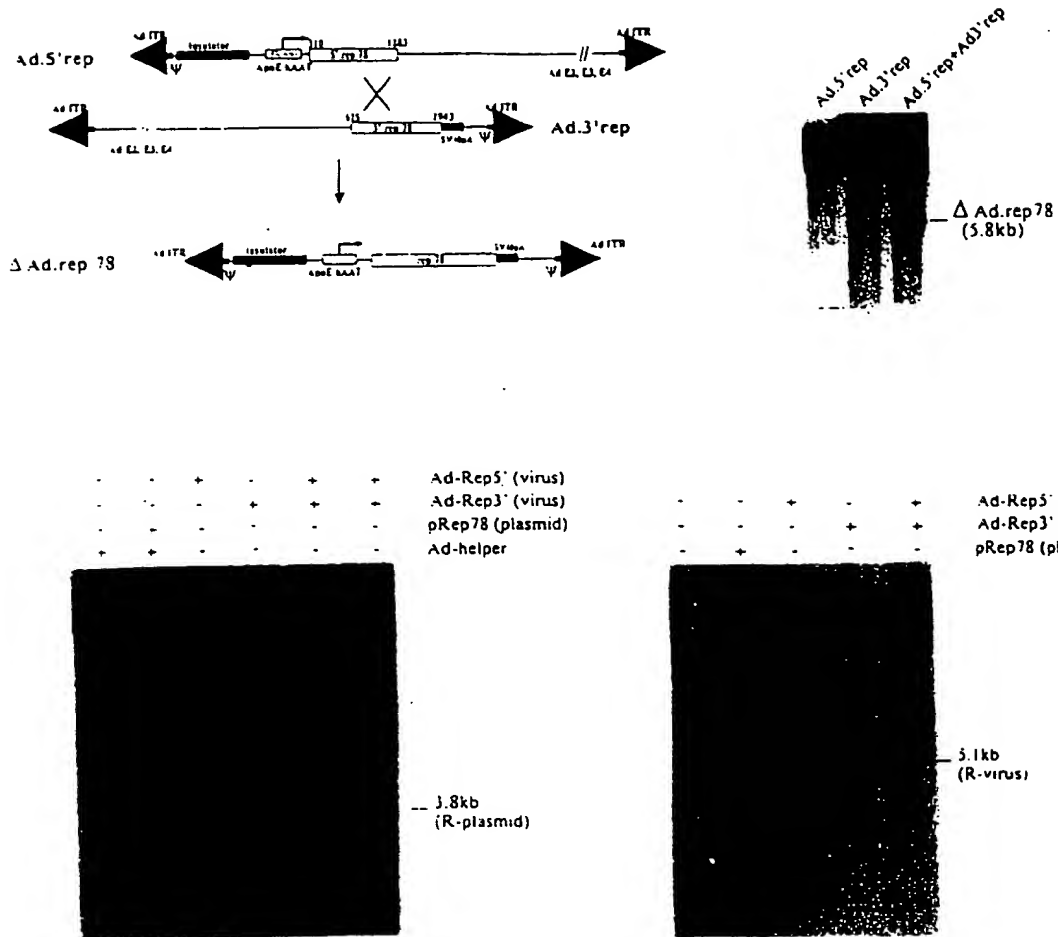


FIG. 29

CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number  
WO 00/73478 A3

(51) International Patent Classification<sup>7</sup>: C12N 15/861,  
15/864, 15/10, A61K 48/00

(21) International Application Number: PCT/US00/15442

(22) International Filing Date: 1 June 2000 (01.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/137,213 1 June 1999 (01.06.1999) US  
60/161,097 22 October 1999 (22.10.1999) US

(71) Applicant (for all designated States except US): UNIVER-  
SITY OF WASHINGTON [US/US]; Suite 200, 1107 N.E.  
45th Street, Seattle, WA 98105-4631 (US).

(72) Inventors; and

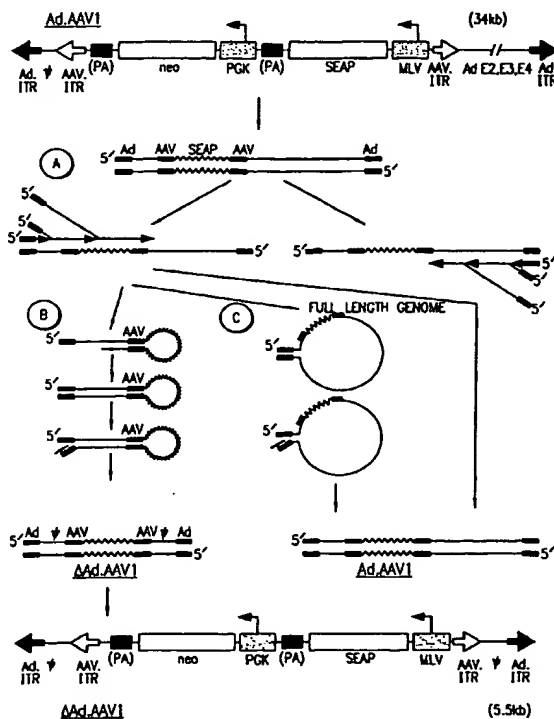
(75) Inventors/Applicants (for US only): LIEBER, An-  
dré [DE/US]; 7251 29th Avenue, Seattle, WA 98115  
(US). SHAYAKHMETOV, Dmitry [RU/US]; 6500 25th  
Avenue, Seattle, WA 98115 (US). FARRAR, Denise  
[US/US]; 1324 W. Emerson Street, Seattle, WA 98119  
(US). PAPAYANNOPOULOU, Thalia [US/US]; 702  
35th Avenue, Seattle, WA 98122 (US).

(74) Agent: ADRIANO, Sarah, B.; Mandel & Adriano, Suite  
60, 35 No. Arroyo Parkway, Pasadena, CA 91103 (US).

(81) Designated States (national): AE, AL, AM, AT, AT (uti-  
lity model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN,  
CR, CU, CZ, CZ (utility model), DE, DE (utility model),  
DK, DK (utility model), DM, EE, EE (utility model), ES,  
FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID,

[Continued on next page]

(54) Title: RECOMBINANT ADENOVIRAL VECTORS EXPRESSING CHIMERIC FIBER PROTEINS FOR CELL SPECIFIC  
INFECTION AND GENOME INTEGRATION



(57) Abstract: The present invention provides for chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_3$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

WO 00/73478 A3



IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report

(88) Date of publication of the international search report:  
5 July 2001

(48) Date of publication of this corrected version:  
24 January 2002

(15) Information about Correction:  
see PCT Gazette No. 04/2002 of 24 January 2002, Section II

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

RECOMBINANT ADENOVIRAL VECTORS EXPRESSING CHIMERIC FIBER PROTEINS FOR CELL SPECIFIC  
INFECTION AND GENOME INTEGRATION5 **FIBER PROTEINS**

This invention was made, at least in part, with funding from the National Institutes of Health (Grant Nos. R01 CA 80192-01 and R21 DK 55590-01). Accordingly, the United  
10 States Government has certain rights in this invention.

**FIELD OF THE INVENTION**

This invention relates to the field of gene therapy, and in particular, to novel adenovirus  
15 (Ad) vectors that selectively infect cells for gene therapy, and to Ad vectors containing modifications of the fiber protein to allow retargeting of any adenovirus serotype.

**BACKGROUND OF THE INVENTION**

20 Gene transfer vectors require the efficient transduction of target cells, stable association with the host genome, and adequate transgene expression in the appropriate target cell, without associated toxic or immunological side effects,. Currently available viral vector systems, including recombinant retroviruses, adenoviruses and adeno-associated viruses, are not suitable for efficient gene transfer into many cell types. Retroviral vectors require  
25 cell division for stable integration. Recombinant adenoviruses are not able to infect many cell types important for gene therapy, including hematopoietic stem cells, monocytes, T- and B-lymphocytes. Moreover, recombinant adeno-associated vectors (AAV) integrate with a low frequency.

30 First generation adenoviruses have a number of properties that make them an attractive vehicle for gene transfer (Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205). These include the ability to produce purified virus at high titers in concert with highly efficient gene transfer of up to 8 kb long expression cassettes into a large variety of cell

- types *in vivo*, including non-dividing cells. Limitations of first generation adenoviruses include the development of immune responses to expressed viral proteins resulting in toxicity and virus clearance. The episomal status of adenoviral DNA within transduced cells is another limitation of first generation Ad vectors. Stable integration of adenovirus DNA into the host genome is reported only for wild-type forms of specific subtypes and appears not to occur in a detectable manner with E1/E3-deleted Ad 5 (adenovirus serotype 5) vectors widely used for gene transfer *in vitro* and *in vivo* [Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205].
- 10 Recombinant AAV vectors (rAAV) integrate with a low frequency (about 1 out of 20,000 genomes) randomly as cocatemers into the host genome (Rutledge, E.A.; Russel, D.W. 1997 *J. Virology*, 71, 8429-8436). The presence of two AAV inverted terminal repeats (ITRs) and as yet unknown host cellular factors seem to be the only requirement for vector integration (Xiao, X., et al, 1997, *J. Virology*, 71, 941-948; Balague, C., et al. 15 1997, *J. Virology*, 71, 3299-3306; Yang, C.C. 1997, *J. Virology*, 71, 9231-9247). In the presence of the large AAV Rep proteins, AAV integrates preferentially into a specific site at human chromosome 19, called AAVS1 (Berns, K.I., 1996, *Fields Virology*, Fields, B.N. et al. (ed) Vol. 2, Lippincott-Raven, Philadelphia, PA, 2173-2220). The AAV capsid is formed by three coat proteins (VP1-3), which interact with specific heparin 20 sulfates on the cell surface and probably with specific receptor(s). However, many cell types, including hematopoietic stem cells, lack these structures so that rAAV vectors based on AAV2 cannot infect or transduce these cells (Malik P. et al., 1997, *J. Virology*, 71, 1776-1783; Quing, K.Y., et al. 1998, *J. Virology*, 72, 1593-1599). Other disadvantages of rAAV vectors include the limited insert size (4.5-5kb) that can be 25 accommodated in rAAV vectors lacking all viral genes and low transducing titers of rAAV preparations.
- Adenovirus infection is initiated by attaching to the cell surface of Ad 5 via its fiber protein (for a review, see Shenk, T. 1996 *Fields Virology*, Vol. 2, Fields, B.N. et al. (ed) 30 Vol. 2, Lippincott-Raven, Philadelphia, PA, 2111-2148). The distal, C-terminal domain of the trimeric fiber molecule terminates in a knob, which binds to a specific cellular

receptor identified recently as the coxsackie-adenovirus receptor (CAR) (Bergelson, J.M. et al. *Science*, 275, 1320-1323). After binding, in an event independent of virus attachment, Arg-Gly-Asp (RGD) motifs in the penton base interact with cellular integrins of the  $\alpha 3$  and  $\beta 5$  types. This interaction triggers cellular internalization whereby the virion achieves localization within the endosome. The endosomal membrane is lysed in a process mediated by the penton base, releasing the contents of the endosome to the cytoplasm. During these processes, the virion is gradually uncoated and the adenoviral DNA is transported to the nucleus where replication takes place. The terminal protein, which is covalently attached to the viral genome and the core protein V that is localized on the surface of the cores have nuclear localization signals (NLSs) (van der Vliet, B. 1995, *The Molecular Repertoire of Adenoviruses*, Vol. 2, Doerfler, W. and Boehm, P.(ed.), Springer Verlag, Berlin, 1-31). These NLSs play a crucial role in directing the adenoviral genome to the nucleus and probably represent the structural elements which allow adenovirus to transduce non-dividing cells. When the double-stranded, linear DNA reaches the nucleus, it binds to the nuclear matrix through its terminal protein.

Since the cell types that can be infected with Ad5 or Ad2 vectors are restricted by the presence of CAR and specific integrins, attempts have been made to widen the tropism of Ad vectors. Genetic modification of adenovirus coat proteins to target novel cell surface receptors have been reported for the fiber (Krasnykh, V. et al. 1998 *J. Virology*, 72, 1844-1852, Krasnykh, V. et al. 1996 *J. Virology*, 70, 6839-6846, Stevenson, S.D., et al. 1997, *J. Virology*, 71, 4782-4790), penton base (Wickham, T.J., et al. 1996, *J. Virology*, 70, 6831-6838; Wickham, T.J., et al. 1995, *Gene Therapy*, 69, 750-756), and hexon proteins (Crompton, J., et al. 1994, *J. Gen. Virol.* 75, 133-139). The most promising modification seems to be the functional modification of the fiber protein or more specifically of the fiber knob as the moiety, which mediates the primary attachment. Two groups have reported the generation of fibers consisting of the Ad5 tail/shaft and the knob domain of Ad3 (Krasnykh, V. et al. 1996 *supra*, Stevenson, S.D., et al. 1997, *supra*). Recently, recombinant adenoviruses with fibers containing C-terminal poly-lysine, gastrin-releasing peptide, somatostatin, E-selectin-binding peptide, or oligo-Histidines were produced in order to change the native tropism of Ad5. Krasnikh et al. found (Krasnykh, V. et al.

1998 *supra*) that heterologous peptide ligands could be inserted into the H1 loop of the fiber knob domain without affecting the biological function of the fiber. Based on studies with other Ad serotypes, it appears that the length of the fiber shaft is a critical element, determining the efficiency of interaction with cell surface integrins and the internalization process. Thus far, there is no reported data demonstrating successful retargeting of Ad5 vectors for a specific cell type.

Therefore, there is a present need for an improved adenovirus vector which can be targeted efficiently to a variety of cell types and tissues and remain stably integrated in the host genome with minimal antigenicity to the host. The present invention discloses novel chimeric adenoviral (Ad) Ad-AAV vectors, which express a modified fiber protein on their capsid, for specifically targeting the vector. Methods of making, uses and advantages of these vectors are described. In addition, the alteration described for the knob and shaft domains of the fiber protein provide a novel approach to retarget any adenovirus serotype for cell specific infection.

### SUMMARY OF THE INVENTION

The present invention provides for novel chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_{3/5}$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

The recombinant adenovirus vectors of the invention (Example I) provide a novel design that allows for the easy production and delivery of a "gutless" adenoviral vector with the added advantage of stable integration of the transgene into the host genome of different cell type. The adenoviral vector described is devoid of all adenoviral sequences except for the 5' and 3' cis elements necessary for replication and virion encapsidation. The adenovirus-associated virus sequences of the invention comprising the 5' (right) and 3' (left) inverted terminal repeats (ITRs) flank the transgene gene cassette such that they



direct homologous recombination during viral replication and viral integration into the host genome. In one embodiment AAV-ITR flanking sequences are used. The vector also contains a selected transgene(s) operably linked to a selected regulatory element and a polydenylation stop signal, which is in turn flanked by the flanking sequences described  
5 above. The selected transgene(s) can be linked under the same regulatory elements or under separate regulatory elements in the same orientation or in opposite orientations with respect to each other. The selected transgene(s) are any gene or genes which are expressed in a host cell or tissue for therapeutic, reporter or selection purposes. This vector is characterized by high titer transgene delivery to a host cell and the ability to  
10 stably integrate the transgene into the host genome. Also provided is a method to improve the integration frequency and site specific integration by incorporating an AAV rep protein into the recombinant hybrid vector.

The invention also provides chimeric fiber proteins (Example II), which includes  
15 naturally occurring fiber proteins in which a portion or portions of the sequence are modified to alter cell or tissue specificity of infection. Altered fiber protein sequences can include fiber protein domains (the knob domain, the shaft domain, and the tail domain) from other or the same adenovirus serotypes or from randomly selected peptides. A chimeric fiber protein can be entirely composed of non-naturally occurring  
20 sequences. The invention further relates to nucleic acid sequences encoding the chimeric fiber proteins. These nucleic acid sequences can be naturally occurring, a mixture of naturally occurring and non-naturally occurring sequences, or entirely non-naturally occurring sequences.

25 The heterologous fiber protein sequences described herein can be inserted into any adenovirus based vector which contains a capsid, rendering the virus capable of specifically infecting a given cell or tissue. Adenoviral vectors having such a heterologous fiber sequence can be used to direct gene transfer into desired cells. For stable integration of the transgene cassette into the host genome, the chimeric Ad-AAV  
30 vector described in the invention is the preferred vector of use.

The invention also includes a library of adenoviruses displaying random peptides in their fiber knobs can be used as ligands to screen for an adenovirus variant with tropism to a particular cell type *in vitro* and *in vivo*.

- 5 The chimeric Ad- vectors described herein include the Ad.AAV genome with a modified fiber protein expressed on its capsid. These chimeric vectors are designed to infect a wide variety of cells, in particular, the cells which can only be poorly transduced by the commonly used retroviral, AAV and adenoviral vectors. These cells include, but not limited to, hematopoietic stem cells, lung epithelial cells, dendritic cells, lymphoblastoid  
10 cells, and endothelial cells. Hematopoietic stem cells such as CD34+ cells can be targeted for gene therapy of sickle cell anemia and thalassemia using the vector described herein. The chimeric Ad-AAV vector capable of transducing genes into endothelial cells can be used in gene therapy for vascular diseases such as atherosclerosis or restinosis after coronary artery surgery.

15

#### **BRIEF DESCRIPTION OF THE FIGURES**

Figs. 1A-1C display a proposed mechanism for forming of  $\Delta$ Ad.AAV1 genome.

- 20 Figs. 2A and 2B show electron photomicrographs of hybrid virus particles: Fig. 2A shows Ad.AAV1 and Fig. 2B shows  $\Delta$ Ad.AAV1.

- Fig. 3 illustrates analysis of  $\Delta$ Ad.AAV1 genomes after transduction of SKHep1 cells. Pulse field gel eletrophoresis (PFGE).  $1 \times 10^6$  control Sk Hep1 cells (SKHep1) (lanes 1-3,  
25 5, 9). SKHep1 cells from G418 resistant pools ( $\Delta$ Ad.AAV1) (infected with  $\Delta$ Ad.AAV1 and selected for 4 weeks) (lanes 6-8, 10-12), or SKHep1 cells collected at 3 days after infection with 2000 genomes Ad.AAV1 (Ad) lanes 4, 13) are sealed in agarose plaques, lysed *in situ* and subjected to PFGE with or without prior digestion with restriction endonucleases. Southern Blot is performed with a SEAP specific probe. U = undigested,  
30 P = digested with PI- Scl, I = I-CeuI, E = EcoRI.



Figs. 4A and 4B show response of K562 and CD34+ cells respectively after infection with  $\Delta$ Ad.AAVBG. Cells are incubated for 6 hours with virus under agitation. At day 3 after infection, transduction frequency is calculated based on the number of X-Gal positive cells. Viability is tested by trypan blue exclusion. N = 3, SEM < 10%.

5

Fig. 5 shows Rep expression in SKHep1 and 293 cells after plasmid transfection.  $5 \times 10^5$  cells are transfected with pAAV/Ad, pRSVrep, or pPGKrep by Ca-phosphate co-precipitation. Three days after transfection, cells are harvested. Lysates are separated on a 10% PA gel, followed by Western Blot with Rep specific antibodies (03-65169),  
10 American Research Products), and developed with ECL (Amersham).

Fig. 6 shows detection of vector integration into AAVS1 by PFGE.

Fig. 7 shows strategy for creating an  $\Delta$ Ad.AAV hybrid vector capable of site-specific  
15 integration. Arrows indicate promoters, (PA) = polyadenylation signal.  $\Psi$  = adenoviral packaging signal.

Fig. 8A-8B shows vectors for transduction studies with SNori as expression unit and analysis of vector integration on genomic DNA from a small cell number. Analogous  
20 vector sets can be generated with  $\beta$ -galactosidase (BG) or green fluorescence protein (GFP) as reporter genes.

Fig. 9 shows strategy for substituting the Ad5 fiber sequence by the heterologous fiber X genes using recombination in *E. coli*.

25

Fig. 10 shows the expression of CAR and  $\alpha_v$ -integrins on test cells. For flow cytometry analysis, HeLa, CHO, K562, and CD34+ cells were incubated with monoclonal anti-CAR (RmcB, 1:400 dilution) or anti- $\alpha_v$ -integrin antibodies (L230, 1:30 dilution). As a negative control, cells were incubated with an irrelevant mouse monoclonal antibody (anti-BrdU,  
30 1:100 dilution). The binding of primary antibodies was developed with anti-mouse IgG-

FITC labeled conjugates (1:100 dilution). Data shown represent the average results of quadruplicate analyses performed on  $10^4$  cells.

Fig. 11 shows the electron microscopy of adenovirus particles. Purified particles from Ad5, 9, and 35 were negative contrast stained and analyzed at a magnification of 85,000x. Defective particles are highlighted by arrows.

Fig. 12 shows the analysis of attachment and internalization of different serotypes to CHO, HeLa, K562, and CD34+ cells. Equal amounts of [ $^3$ H]-thymidine-labeled virions of Ads 3, 4, 5, 9, 35, and 41 (measured by OD<sub>260</sub>, and equivalent to an MOI of 400 pfu per cell for Ad5) were incubated for one hour on ice as described in Materials and Methods. Cells were then washed, and the number of labeled virions bound per cell was determined. For internalization studies, viruses were first allowed to attach to cells for 1 h on ice. Then, unbound viral particles were washed out. Cells were then incubated at 37°C for 30 min followed by treatment with trypsin-EDTA and washing to remove uninternalized viral particles. The data were obtained from two to four independent experiments performed in triplicate. Note the different scale on the Y-axes for CD34+ cells.

Figs. 13A-13C show attachment and internalization of different adenovirus serotypes to HeLa, CHO and 293 cells respectively.

Figs. 14A and 14B show attachment and internalization of different adenovirus serotypes to CD34+ and K-562 cells respectively.

Fig. 15A-15C shows the analysis of viral replication in K562 and CD34+ cells by Southern blot analysis of methylated viral DNA. Replication studies were performed with  $1 \times 10^5$  K562 cells (A) or CD34+ cells (B), infected with methylated Ad5, Ad9 or Ad35. The lane labeled as "load" represents DNA that was extracted from the media/cell mixture immediately after adding the indicated viral dose to cells. The intensities of bands corresponding to methylated and un-methylated viral DNA indicate that ~85% of

the input virus was methylated. To quantify adsorption and internalization, DNA analysis was performed after prior incubation of virus with cells at 0°C (adsorption) or 37°C (internalization). For dose dependent replication studies, the indicated viral dose (expressed as the number of genomes) was added to the cells, and cellular genomic DNA together with viral DNA was extracted 16 hours or 36 hours post-infection for K562 and CD34+ cells, respectively. Identical amounts of sample DNA were analyzed by Southern blot. For quantification purposes, Ad9 replication was analyzed together with Ad5 using an Ad5/9 chimeric probe that hybridizes with DNA of both serotypes (C). The analysis of Ad5 versus Ad35 replication was performed with the corresponding Ad5/35 chimeric probe. Since separate hybridizations with both Ad5/35 and Ad5/9 probes gave identical signal intensities for Ad5 DNA only one panel is shown for Ad5 replication in test cells. To produce distinguishable fragments specific for the methylated or non-methylated status of viral genomes, Ad5 DNA was digested with *Xho* I, while Ad9 and Ad35 DNA was digested with *Xho* I and *Hind* III. The bands specific for methylated (not-replicated) viral DNA were ~12kb for Ad9, 35kb for Ad5, and ~12kb for Ad35. The fragments specific for non-methylated DNA were 5.8kb for Ad9, 6.1kb for Ad5, and 9.5kb for Ad35. Chimeric Ad5/9 and Ad5/35 DNA fragments (1.8kb) were used as quantification standards and applied onto gel together with digested viral/cellular DNA (shown on the left part of the figures).

Fig. 16A-16B shows the structure of Ad5GFP and chimeric Ad5GFP/F35 vectors. A) Schematic diagram of the original E1/E3 deleted Ad5-based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes. A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The

Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal. The region of Ad5GFP/F35 encoding for chimeric fiber was completely sequenced with Ad5 specific primers (see Material and Methods). B) Restriction analysis of viral genomes. Viral DNA was isolated from purified Ad5GFP and Ad5GFP/F35 particles as described elsewhere. One microgram of DNA was digested with *Hind* III or *Kpn* I and separated in ethidium bromide stained agarose gels (left panel) which were subsequently blotted and analyzed by Southern blot with an Ad5 E4 specific probe (nt 32,7775-33,651) (right panel). Specific patterns, designating the correct structure for both viral vectors were detected. The *Hind* III fragments specific for Ad5GFP and Ad5GFP/F35 were 2.9kb and 4.9kb, respectively. The *Kpn* I fragment that confirmed the correct Ad5GFP/F35 structure was 1.6kb compared to a 7.6kb Ad5GFP fragment. M - 1kb ladder (Gibco-BRL, Grand Island, NY).

Fig. 17 shows the generation of  $\Delta$ Ad.AAV genomes by recombination between inverted homology regions. Recombination between two inverted repeats (IR) present in one Ad.AAV vector. The first-generation Ad.AAV vector (~34kb) contains two 1.2kb inverted homology elements flanking the transgene cassette. One AAV-ITR is inserted between the Ad packaging signal ( $\psi$ ) and the left IR. During Ad replication, recombination between the Irs mediates the formation of the  $\Delta$ Ad.AAV genomes with the transgene flanked by Irs, AAVITRs, Ad packaging signals, and Ad ITRs. These genomes are efficiently packaged into Ad capsids.

Fig. 18 shows the structure of Ad5/11, Ad5/35. Schematic diagram of the original E1/E3 deleted Ad50based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes.

A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal.

5 Fig. 19 shows the cross-competition for attachment and internalization of labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled viruses, and with anti-CAR or anti- $\alpha_v$ -integrins Mab. (A) For attachment studies,  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled competitor virus at 4°C for 1 h. Then, equal amounts of [ $^3$ H]-labeled viruses, at a dose equivalent to an MOI of 100 pfu per cell  
10 determined for Ad5GFP, were added to cells followed by incubation at 4°C for 1 h. Cells were then washed with ice-cold PBS, pelleted and the percentage of attached virus (cell-associated counts per minute) was determined. For analysis of cross-competition for internalization, cells were pre-incubated with a 100-fold excess of competitor virus at 37°C for 30 min before labeled virus was added. After an additional incubation at 37°C  
15 for 30 min, cells were treated with trypsin-EDTA for 5 min at 37°C, washed with ice-cold PBS, pelleted, and the percentage of internalized virus was determined. For controls, cells were incubated with labeled viruses without any competitors. Preliminary experiments had shown that the conditions chosen for competition studies allowed for saturation in attachment/internalization on K562 cells for all unlabeled competitors. (B)  $10^5$  K562 cells  
20 were pre-incubated for 1 hour at 4°C with anti-CAR MAb (RmcB, diluted 1:100) or with anti- $\alpha_v$ -integrin MAb (L230, diluted 1:30), followed by incubation with labeled viruses according to the protocols for attachment or for internalization as described above. For each particular serotype, the percentage of attached/internalized virus was compared to the control settings, where cells were preincubated under the same conditions with a  
25 1:100 dilution of an irrelevant antibody (anti-BrdU Mab) before addition of the labeled virus. Note that the specific competitors but not the corresponding controls significantly inhibited Ad5 internalization to a degree that is in agreement with published data (59).  
N>/=4

30 Fig. 20. Cross-competition for attachment and internalization of [ $^3$ H]-labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled Ad3 virus (A), and of [ $^3$ H]-



labeled Ad3 virions with unlabeled viruses (B).  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled viral particles according to attachment or internalization protocols described for Fig.6. Equal amounts of [ $^3$ H]-labeled Ad5GFP, Ad5GFP/F35, or Ad35 (A) or [ $^3$ H]-labeled Ad3 (B) were added to cells at a dose equivalent to an MOI of 100 pfu per cell for Ad5GFP. In control settings, cells were incubated with labeled viruses without any competitors. N=4.

Fig. 21 shows the transduction of CD34+, K562, and HeLa cells with Ad5GFP and chimeric Ad5GFP/F35 vectors.  $1 \times 10^5$  cells were infected with different MOIs (pfu/cell) of viruses in 100  $\mu$ l of media for 6 hours at 37°C. Virus containing media was then removed, and the cells were resuspended in fresh media followed by incubation for 18 h at 37°C. The percentage of GFP expressing cells was determined by flow cytometry. N=3

Fig. 22 shows the distribution of GFP-positive cells in subpopulations of human CD34+ cells expressing CAR or  $\alpha_v$ -integrins.  $1 \times 10^5$  CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu/cell as described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CAR (1:100 final dilution) or anti- $\alpha_v$ -integrin (1:30 final dilution) primary MAbs for 1 h at 37°C. Binding of primary antibodies was developed with anti-mouse IgG-PE labeled secondary MAbs (1:100 final dilution) at 4°C for 30 min. For each variant,  $10^4$  cells were analyzed by flow-cytometry. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The data shown were representative for three independent experiments.

Fig. 23A-23B shows the distribution of GFP-positive cells in a subpopulation of human CD34+ cells, expressing CD34 and CD117 (c-kit). (A) Co-localization of GFP expression with CD34 or CD117: CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell under the conditions described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CD34 PE-conjugated MAbs (final dilution 1:2) or with anti-CD117 PE-conjugated MAbs (final dilution 1:5)

for 30 min on ice, and  $10^4$  cells per variant were subjected to two-color flow cytometry analysis. For negative control staining, no antibodies were added to the cells before analysis. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The experiment was performed two times in triplicates, and typically obtained results are shown. The SEM was less than 10% of the statistical average. (B) Transduction of CD34+/CD117+ cells with Ad5GFP and chimeric Ad5GFP/F35 virus vectors: CD34+ cells, cultured overnight before staining in media without SCF, were incubated with PE-labeled anti-CD117 MAb for 30 min on ice. The fraction of CD117-positive cells was sorted by FACS. More than 97% of sorted cells were positive for CD117.  $1 \times 10^5$  CD117+/CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell, as for Fig.8. Twenty-four hours post infection, the percentage of GFP positive was determined by flow cytometry. For mock infection, CD117+/CD34+ cells were incubated with virus dilution buffer only. The infections were done in triplicates, and the average percentage of GFP-expressing cells is indicated on the corresponding histogram. The SEM was less than 10% of the statistical average.

Fig. 24 shows the southern analysis of viral genomes in GFP-positive and GFP-negative fractions of CD34+ cells infected with the Ad5GFP and chimeric Ad5GFP/F35 vectors. CD34+ cells were infected with viruses at an MOI of 100 as described for Fig.21. Twenty four-hours post infection, cells were sorted by FACS for GFP positive and GFP negative fractions.  $10^5$  cells from each fraction were used to isolate genomic DNA together with viral DNA. Before cell lysis, a rigorous treatment with trypsin and DNase followed by washing was performed to exclude that genomic DNA samples were contaminated by extracellular viral DNA. A) The upper panel shows the ethidium bromide stained 1% agarose gel before blotting demonstrating that similar amounts of genomic DNA were loaded. This amount corresponded to DNA isolated from ~25,000 GFP+ or GFP- cells. The lane labeled Aload@ represents viral DNA purified from Ad5GFP or Ad5GFP/F35 virions mixed with pBluescript plasmid DNA (Stratagene) as a carrier and applied on a

gel at the amount that was actually used to infect 25,000 cells. As a concentration standard, a serial dilution of Ad5GFP genomes was loaded on the gel (left side). For Southern analysis (lower panel), an 8kb-long HindIII fragment corresponding to the E2 region of Ad5 was used as a labeled probe. Hybridized filters were subjected to  
5 PhosphoImager analysis and then exposed to Kodak-X-OMAT film for 48 h at B70°C. The cellular/viral genomic DNA is indicated by an arrow. (B) To detect Ad5GFP genomes in transduced cells, PCR amplification followed by Southern blot hybridization was performed on the same samples that were used for quantitative Southern blot hybridization in (A). DNA purified from ~2,500 cells was subjected to PCR (95°C B1min,  
10 53°C-1min, 72°C B1min, 20 cycles with primers Ad5-F1 and Ad5-R1). One fifth of the PCR reaction was subjected to agarose gel electrophoresis (upper panel). A 0.9 kb-long DNA fragment, specific to the E4 region of Ad5 was detected for transduced Ad5GFP/F35 genomes. DNA then was blotted onto Nybond-N+ membrane and Southern blot hybridization (lower panel) with an Ad5 E4 specific DNA probe was performed. In  
15 addition to the 0.9kb DNA fragment, the PCR primers generated a smaller 0.5 kb-long fragment that also hybridized with with the E4 region probe.

Fig. 25 shows the role of fiber shaft length in Ad infection strategies. CAR binding (Ad5 and Ad9) variants and Ad35, which interacts with a non-CAR receptor were analyzed on  
20 CAR expressing cells (293, Y79) and K562 cells which do not express significant CAR amounts. All vectors contain a GFP expression cassette packaged into an Ad5 capsid with modified fibers.

Fig. 26 shows the tertiary structure of Ad5 knob: localization of CAR binding sites, H-I  
25 loop and G-H loop.

Fig. 27 shows the substitution of the G-H loop with heterologous peptides.

Fig. 28 shows the attachment and internalization of metabolically labeled serotypes with  
30 human cell lines.



within a host cell. Accordingly, the transgene sequence may be a replicon sequence that directs replication of the vector within the host cell, resulting in an autonomously replicating vector. Alternatively, the transgene sequence may permit vector replication that is dependent upon the host cell's replication machinery.

5

The base vector sequences may be a operatively linked to a transgene sequence that encodes a gene product, such as a polypeptide, rRNA, or tRNA. For example, the transgene may encode a polypeptide such as a viral capsid protein, or a viral fiber protein. The transgene may be derived from the same or different serotype as the base vector sequence.

10

Another example of a transgene includes a reporter gene that encodes a gene product that can be used as a selectable marker, such as drug resistance or a colorimetric marker. The reporter gene may encode a gene product which can be readily detected by, for example, a visual microscopic, immunochemical, or enzymatic assay. The preferred reporter gene encodes a gene product that can be detected by a non-destructive method that does not destroy the cell that expresses the reporter gene.

15

A therapeutic gene is another example of a transgene. A therapeutic gene encodes a gene product (e.g., polypeptide or RNA) which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or the tissue or the organ or the organism containing the host cell. The therapeutic benefit may result from modifying a function of a gene in the host genome or from the additional function provided by the therapeutic protein, polypeptide or RNA.

20

The base vector sequence may be linked to a transgene sequence that is an regulatory element, such as a promoters, enhancers, transcription termination signals, polyadenylation sequences. The regulatory element may direct expression of the transgene sequence that encodes a gene product by direct transcription or translation.

25

The regulatory element may regulate the amount or timing of expression of the transgene

30

sequence. The regulatory element may direct expression of the transgene in certain host cells or tissues (e.g., host-specific or tissue-specific expression).

5 The base vector sequence may be linked to a transgene sequence that permits the vector, to integrate into another nucleotide sequence. The integration sequence may direct integration of the whole vector or portions of the vector. The integration sequence may or may not be related to the base vector sequence. For example the integration and base vector sequences may be from the same or different viral serotype. The integration sequence may be inverted repeat sequences (ITRs) from adenovirus (Ad), adenovirus-associated virus (AAV), or HIV.  
10

The base vector sequence may be linked to a transgene sequence that directs homologous recombination of the vector into the genome of a host cell. Such transgene sequences may or may not be from the same viral serotype as the base vector sequence.  
15

The vector may be used to transport the heterologous sequence into a host cell or into a host cell's genome.

The vector may comprise multiple endonuclease restriction sites that enable convenient insertion of exogenous DNA sequences.  
20

The term "hybrid vector" as used in the invention refers to a vector which comprises a nucleic acid sequence combined from two different viruses (e.g. Adenovirus and AAV).

25 "Chimeric vector" refers to a vector which contains nucleic acid sequences that are unnatural to the base vector (i.e. sequences not occurring naturally or sequences not in their natural background including heterologous sequences). A chimeric vector as used in the invention may also be a hybrid vector. An example of a chimeric vector is Ad.AAV expressing a modified fiber protein on its capsid.

30

The term "transduction" or "infection" refers to a method of introducing viral DNA within a virus particle into a host cell. The viral DNA herein is in the form of recombinant virus, which is generated by linking a segment of DNA of interest into the viral genome in such a way that the gene can be expressed as a functional protein.

5

The term "transfection" refers to a method of introducing a DNA fragment into a host cell.

10 The term "heterologous" as used herein means that a nucleic acid or peptide sequence is placed in a context that is not endogenous to the base adenovirus vector or to a transduced cell. For example, a peptide sequence can be transferred from a protein to another protein, the resulting protein is referred to herein as heterologous protein. A chimeric fiber protein, (e.g., a serotype 5 tail domain and a serotype 35 shaft and knob domain) is considered a "heterologous" to the Ad 5 vector. The term also includes  
15 nucleic acids (e.g. coding sequences) from one strain or serotype of adenovirus introduced into a different strain or serotype of adenovirus.

The term "regulatory elements" is intended to include promoters, enhancers, transcription termination signals, polyadenylation sequences, and other expression control sequences.  
20 Regulatory elements referred to in the invention include but are not limited to, those which direct expression of nucleic acid sequence only in certain host cells (e.g. tissue specific regulatory sequences).

The term "operably linked" indicates that a polynucleotide sequence (e.g. a coding  
25 sequence or gene) is linked to a regulatory element in such a way that the regulatory element sequence controls and regulates the transcription or translation or both of that polynucleotide sequence. The orientation of the regulatory element may vary (eg, be in reverse orientation with respect to the right ITR). The term also includes having an appropriate start signal (e.g., ATG) in front of the polynucleotide sequence to be  
30 expressed and maintaining the correct reading frame to permit expression of the polynucleotide sequence under the control of the expression control sequence and

production of the desired polypeptide or protein. Regulatory sequences can also include 3' sequences which ensure correct termination (eg. polyadenylation stop signal).

5 The term "gene therapy" used herein, refers to a method which introduces a segment of exogenous nucleic acid into a cell in such a way that it results in functional modification to the recipient cell by expression of the exogenous nucleic acid. The exogenous nucleic acid is typically therapeutic in that the expression of the encoded protein, polypeptide or RNA corrects cellular dysfunction due to a genetic error or more generally counteracts any undesirable functions which are associated with a genetic or acquired disease. The  
10 term "exogenous nucleic acid" refers to DNA or RNA sequences not normally expressed in the treated transformed cell. The term also refers to DNA and RNA sequences which are expressed in a treated transformed cell at a higher, lower or in an otherwise different pattern than in the untreated, nontransformed cell. This non-natural expression can also be termed heterologous expression.

15 A "gene therapy vector" refers to a vector used for gene therapy. i.e. to introduce the exogenous nucleic acid into a recipient or host cell. The exogenous nucleic acid may be transiently expressed or integrated and stably expressed in the recipient or host cell.

20 The term "plasmid" as used herein refers to any nucleic acid molecule which replicates independently of the host, maintains a high copy number, and which can be used as a cloning tool.

The term "parallel strand of DNA" and "anti-parallel strand of DNA" refers to as each of  
25 the strands of DNA of the double stranded adenovirus. The Figures diagram the location of certain nucleotides on the parallel strand of DNA. The anti-parallel strand of DNA refers to the other of the two strands of DNA which is not depicted in the Figures. The fiber protein is encoded on the anti-parallel strand of DNA. To simplify the vector diagrams, the fiber sequences are shown on the parallel strand even though the gene is  
30 located on the anti-parallel strand.



The term "reporter gene" refers to any nucleic acid sequence which encodes a polypeptide or protein which can be readily detected by, for example, a visual, microscopic, immunochemical or enzymatic assay. Preferred reporter genes are those  
5 that can be detected by a non-destructive method that does not destroy the treated, transformed cells or tissue.

The term "selection gene" used herein refers to any nucleic acid fragment which encodes a polypeptide or protein whose expression is used to mark a cell as a transformed cell by  
10 a given vector.

The term "therapeutic gene" refers herein to a DNA fragment encoding a functional polypeptide, protein or RNA, which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or to the organ or organism containing the host  
15 cell. The therapeutic benefit may result from modification of a function of a native gene in a host or from the additional function provided by the therapeutic protein, polypeptide or RNA.

The term "host tissue" or "host cell" as used herein, refers to a tissue or cell in which a  
20 therapeutic gene is to be expressed to modify its function.

It is well-known in the biological arts that certain amino acid substitutions may be made in protein sequences without affecting the function of the protein. Generally, conservative amino acid substitutions or substitutions of similar amino acids are tolerated  
25 without affecting protein function. Similar amino acids can be those that are similar in size and/or charge properties, for example, aspartate and glutamate, and isoleucine and valine, are both pairs of similar amino acids. Similarity between amino acid pairs has been assessed in the art in a number of ways. For example, Dayhoff et al. (1978) in *Atlas of Protein Sequence and Structure*, Volume 5, Supplement 3, Chapter 22, pp. 345-352,  
30 which is incorporated by reference herein, provides frequency tables for amino acid substitutions which can be employed as a measure of amino acid similarity. Dayhoff et

al's frequency tables are based on comparisons of amino acid sequences for proteins having the same function from a variety of evolutionarily different sources. Therefore, any obvious changes in the amino acid sequences (as described above) to the sequences of the invention are already contemplated.

5

Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups include but are not limited to: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, asparagine-glutamine, and aspartate-glutamate. Therefore, polypeptide substitution for "substantially similar" sequences (as described above) to the amino acid sequences described invention are already contemplated.

20

In order that the invention herein described may be more fully understood, the following description is set forth.

The present invention provides unique gene transfer vehicles which overcome many of the limitations of prior art vectors. The invention describes a first generation adenovirus vectors comprising left and right Ad ITRs, an Ad packaging sequence, a transgene cassette with regulatory elements, and a pair of cassette ITRs flanking the transgene cassette that direct predictable viral genomic rearrangements during viral replication as well as direct the integration of the transgene cassette into the host cell genome.

30

One predictable rearrangement that occurs during viral replication is the generation of a gutless adenovirus vector (also referred to herein as  $\Delta$ Ad) that comprises right and left Ad ITRs, an Ad packaging sequence, a transgene cassette flanked by cassette ITRs and the gutless vector is devoid of all other immunogenic viral genes.

5

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the transgene cassette is directed by co-infection with an Ad vector expressing e.g., the rep 78 protein to achieve site-specific integration in the e.g., AAVS1 site on human chromosome 19.

10

The invention further describes a novel way of targeting these recombinant adenovirus vectors to selected cells by modifying the adenovirus fiber protein that is expressed on the capsid. Changes to both the fiber shaft and the fiber knob domain proved to successfully retarget the Ad vector to a desired cell type. In addition, the G-H loop within the fiber knob domain is identified as a novel site that affect the binding affinity and specificity of the recombinant adenovirus vector. Substitution of peptide sequences into the G-H loop retarget the gutless vector to a desired cell type.

15

An adenovirus display library has been generated that expresses random peptides within the G-H loop of the fiber protein. This type of a library is used as ligands to screen for adenovirus vectors that bind to desired cell types. One advantage of using an adenovirus display library versus a phage display library is that once adenovirus affinity to a desired cell is identified the targeted adenovirus vector is ready to accept a transgene cassette and can be used to generate a gutless adenovirus vector, for example for use in gene therapy.

20

The chimeric Ad vectors described below contain a modified fiber protein in the capsid of the adenovirus which renders the vector capable of infecting a desired cell types. Therefore, according to the invention, a gutless chimeric  $\Delta$ Ad-AAV vector can be generated to introduce any transgene(s) into any host cell or tissue which is normally refractory to most commonly used gene therapy viral vectors. In addition, the chimeric  $\Delta$ Ad-AAV vector of the invention, is devoid of adenoviral genes, and contains AAV ITR

25

30

sequences that flank the transgene cassette, which direct stable transgene integration in the host genome allowing long term expression of the transgene.

The transgene cassette described in the invention may carry a transgene which is either a reporter gene, a selectable gene for in vitro or in vivo selection of transduced cells, or a therapeutic gene. In one embodiment of the invention the reporter transgene can be but is not limited to,  $\beta$ galactosidase. Many reporter genes are commonly used in the art, of which any could be carried as a transgene in the Ad.AAV vector of the invention. Other examples of reporter genes are GFP and alkaline phosphatase.

The following describes an embodiment of the first generation Ad vectors of the invention having a wild-type capsid and a transgene cassette flanked by cassette ITR sequences; (b) fiber protein that is modified to retarget Ad vectors; and (c) the combination of both technologies that enables the production of chimeric  $\Delta$ Ad vector including a modified fiber protein expressed on the capsid which retargets the base vector to a desired cell type for infection and transgene integration.

A. Integrating Ad hybrid vectors of the invention:

It has been shown that inverted repeats (IRs) inserted into the E1 region of AdE1- vectors can mediate predictable genomic rearrangements resulting in a gutless vector genome devoid of all viral genes. A specific embodiment of such IR-mediated rearrangements is the Adeno-AAV, first generation hybrid adenovirus vector containing AAV inverted terminal repeats (ITR) flanking a transgene cassette. The AAV ITRs mediate the formation of a genome similar to that of the  $\Delta$ Ad.IR genome (Steinwaerder et al., 2000 Journal of Virology).  $\Delta$ Ad vectors devoid of all viral genes stably integrate and transduce cultured cells with efficiencies comparable to e.g. rAAV vectors. The Examples demonstrate by Southern blot analysis that the  $\Delta$ Ad vectors integrate randomly into the host genome.

The Ad vectors of the invention comprise a left Ad ITR, an adenovirus packaging sequence located 3' to the Ad ITR; a transgene cassette located 3' to the packaging

sequence comprising a polyadenylation signal, a transgene, and a heterologous promoter, and flanked by a pair of cassette ITRs. Adenoviral genes used for replication such as E1, E2, E3, E4 are located 3' to the right cassette ITR and a right Ad ITR is located 3' to the replication genes. The vectors of the invention are particularly suited to treat: genetic disorders, cancers, and infectious diseases (such as HIV, emboli, or malaria). Treatable genetic diseases such as hemophilia A and B; cystic fibrosis; muscular dystrophy, and  $\alpha_1$ -antitrypsin disorder are ideal candidates for genetic disease that can be treated by vectors of the invention. A specific example of a therapeutic gene to combat a genetic disorder is gamma-globin to ameliorate sickle cell anemia.

To aid in the selection of transduced cells and characterize the integration site of the transgene cassette, an embodiment of the invention includes the addition of a sequence comprising a bacterial origin of replication, plus a selectable gene. An embodiment of this is an SNori sequence added to the transgene cassette. This allows the  $\Delta$ Ad to be expressed in human and bacterial cells, therefore allowing selection of the transduced cells and characterization of the integration site in the genome of transduced mammalian cells.

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein in 293 cells to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration, which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

Integration of the transgene cassette contained in the adenoviral vectors into chromosomes may be associated with silencing (or blocking) of transgene expression. The silencing of transgenes can be overcome by adding insulator elements to the transgene cassette. For example, HS-4 insulator elements derived from the chicken-globin LCR can function in Ad vectors to shield heterologous promoters from adenoviral enhancers. HS-4 insulators or the *Drosophila* Gypsy gene can also be used to prevent silencing transgenes.

Another embodiment of the invention is to split the transgene cassette into two portions of the transgene each carried in a different recombinant adenoviral vector of the invention. Each portion of the same transgene has an overlapping region of homology. After infection with both vectors, each carrying the different but overlapping portion of the same transgene, homologous recombination event occurs resulting in the reconstitution of the complete transgene which is then expressed. This technique is used to produce hybrid adenoviral vectors that accommodate large inserts including, but not limited to a 13kb genomic hAAT gene or a 12kb  $\gamma$ globin LCR $\gamma$ globin expression cassettes for ameliorating sickle cell anemia (or correcting  $\gamma$ -globin mutations). The formation of the hybrid  $\Delta$ Ad vector genomes, after recombination between two vectors, is more efficient if the overlapping region of homology within the transgene is longer.

An advantage of the present invention is a method to rapidly isolate pure gutless hybrid adenoviral vectors such as  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV<sup>fx</sup> vectors. To minimize the contamination of  $\Delta$ Ad with first generation vectors (Ad vectors) a strategy is described in Example I H. It is anticipated that these approaches will yield the same titer of  $\Delta$ Ad vectors, however the contamination with full-length genome vectors will be less. This improved isolation of the vectors is extremely important to avoid toxic side effects after in vivo application.

**B. Tropism modified adenovirus vectors:**

The Ad vectors of the invention can be modified so that they target a host cell of interest. There are more than 50 human Ad serotypes (Appendix I), including variants with different tissue selectivity or tropism. It is accepted in the art that different Ad serotypes bind to different cellular receptors and use different entry mechanisms. Most recombinant  
5 adenovirus vectors use adenovirus serotype 5 as the base vector serotype 5 (Ad5) (Hitt, M.M., et al, 1997, Adv. in Pharmacology 40, 137-205). Ad5 infection is primarily mediated by its fiber protein binding to CAR and secondarily by its penton base protein binding to integrin. Due to the lack of CAR and/or integrin expression on many cell and tissue types, Ad5 mediated gene transfer is inefficient in a number of tissues which are  
10 important targets for gene therapy such as endothelia, smooth muscle, skin epithelia, differentiated airway epithelia, brain tissue, peripheral blood cells, or bone marrow. The following describes Ad5 vectors of the invention having a change in infectivity and tropism as a result of altering the fiber protein sequence.

15 The infectivity of different Ad serotypes is limited to a number of human cell lines. Infectivity studies revealed that Ad5 and Ad3 are particularly suitable for infecting and targeting endothelial or lymphoid cells, whereas Ad9, Ad11 and Ad35 efficiently infected human bone marrow cells. Therefore, the knob domain of the fiber protein of Ad9, Ad11 and Ad35 are excellent candidates for retargeting the Ad5 vector to human bone marrow  
20 cells. Other possible serotypes include Ad7.

In the modified fiber protein of the invention the fiber knob domain of the Ad5 fiber has been replaced with another Ad serotype fiber knob domain. An embodiment of the invention is the modified Ad5/35 fiber protein (a recombinant Ad5 vector expressing a  
25 modified fiber protein comprising of a fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad35). The Ad5/35 chimeric fiber protein shows a broader spectrum of infection to a subset of CD34+ cells, including those with stem cell activity. The Ad5/11 chimeric fiber protein (a recombinant Ad5 vector expressing a modified fiber protein comprising the fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad11)  
30 showed similar tropism.

In addition to the knob domain modifications, the invention describes the added advantage of modifying both the fiber shaft domain and the fiber knob domain to produce a shortened fiber protein. The length of the fiber shaft domain plays a key role in the host receptors used for viral vector entry into the host cell. To show this Ad5, Ad5/9, and  
5 Ad5/35 variants were constructed with long (22 $\beta$ -sheets) and short-shafted (7 $\beta$ -sheets)-shafted fibers. These analyses demonstrated that efficient viral infection involving CAR as the primary receptor for Ad5, Ad5/9 requires a long-shafted fiber protein, whereas the cell entry strategy of Ad5/35 (which binds to an still uncharacterized non-CAR receptor) does not depend on the shaft length (Fig. 3). The modification in both the fiber shaft  
10 domain length (between 5 > 10  $\beta$ -sheets) and the fiber knob domain (from a different Adserotype than the base vector is a novel mode of altering Ad vector tropism.

To broaden the repertoire of cell types that Ad vectors can infect, a specific binding region, the G-H loop, within the knob domain has been newly identified herein to  
15 improve binding affinity and specificity. Alteration within this region will redirect the Ad vector to a desired cell type. For example, the invention describes the G-H loop sequence within the fiber protein knob domain, which can be replaced with heterologous peptide ligand sequences without affecting the functionally important tertiary structure of the Ad fiber knob domain, while changing the binding affinity and specificity of the vector (Figs.  
20 6,7). This G-H loop region is exposed on the central part of the knob surface and may be strategically a better site for incorporation of heterologous ligands than the peripheral H-I loop (Krasnykh, V. et al., 1998, *J. Virol.*, 72:1844-52.) of the knob C-terminus (Michael, S. I., et al., 1995, *Gene Ther.*, 2:660-8., Wickham, T. J. et al., 1996, *Nat. Biotechnol.*, 14:1570-3.), which are the substitution sites used by others. Therefore, these G-H loop  
25 modifications within the fiber knob domain will allow the Ad vector to be redirected to infect a desired cell type, as long as the G-H loop ligand sequence binds to at least one surface protein on the desired cell type. Fig. 7 shows some possible substitutions. Example II J demonstrates that the virion tolerates the insertion of a cycling peptide (12 amino acids) with a constrained secondary structure that allows the exposure on the knob  
30 surface. A defined ligand (RGD) can be inserted into the G-H and the H-I loop of an



Ad5 capsid that is ablated for CAR, and integrin tropism. Infectivity studies show the potential advantage of this new insertion site.

Use of the Vectors of the Invention for "Gene Therapy"

5

The liver is the major organ for protein synthesis. Therefore an important goal of gene therapy is to target gene therapy vectors to the liver. To genetically correct many types of mutant proteins hepatocytes need to be infected with gene therapy vectors carrying a corrected transgene. Example II J describes a G-H loop substitution in the knob domain  
10 of the fiber protein with both RI and RII+ (of the malaria circumsporozoite surface protein) in a short shafted fiber protein which directs the vector to have affinity and specificity to hepatocytes.

Example II K applies a similar protocol to alter the fiber knob domain in the G-H loop  
15 region with peptides that target the vector breast cancer cell lines (MDA-MB-435). These novel approaches to redirect vectors described in the invention allow lower doses of the gene therapy vectors to be administered with a higher safety profile.

In example II L a protocol for preparing an adenovirus display library is described that  
20 uses the fiber knob protein to display a library of random peptide sequences within the G-H loop. This library of adenoviruses with modified fiber proteins is screened for affinity and specificity for a desired cell type. There are two main advantages of using this adenovirus display library to screen for target peptides that allow binding to a desired cell type over a phage display library system. First, once a ligand peptide is identified that  
25 binds to the desired cell type it is already in the vector of choice for gene therapy delivery. The peptide does not need to be engineered into another vector, as is the case for the phage display library vectors. This reduces the steps required to identify a targeted fiber protein for a desired cell type. The second advantage of this method, is that the adenoviruses are able to display multiple copies of the modified fiber protein on their  
30 capsid. This allows for dimerization and trimerization of the fiber protein with the host cell receptor. The multimerization of fibers proteins is a realistic, in vivo interaction of

the trimeric fiber protein with the host cell receptor. In contrast, phage vectors can only display one fiber peptide sequence on their surface, which significantly limits the ability of interaction with host cell surface receptors.

5 C. A chimeric adenovirus vector with selective tropism:

The chimeric vectors of the invention combine two vectors: an Ad. ITR and a Ad. fx where fx describes a modified fiber protein. A first generation adenovirus vector of serotype 5 is the base vector that carries a transgene cassette flanked by heterologous  
10 ITRs. These specific inverted terminal repeat sequences, such as AAV ITRs direct stable integration of the transgene cassette into the host genome as well as control predictable genomic rearrangements that occur during viral replication. This vector can also carry a modified fiber gene (described in Examples II). During replication predictable genomic rearrangements occur which result in the generation of a gutless adenovirus vector (e.g.  
15  $\Delta$ Ad.AAV<sup>fx</sup>) which expresses the modified fiber protein on its capsid. The modified fiber protein allows the gutless vector to be targeted to a selected cell type. The targeted vector is a gutless adenovirus vector devoid of adenoviral genes which can integrate its transgene into the host genome. The transgene cassette can carry reporter, selectable, or therapeutic genes.

20

In one embodiment of the invention, the gutless targeted  $\Delta$ Ad.AAV<sup>fx</sup> carries the reporter gene of  $\beta$ galactosidase ( $\Delta$ Ad.AAV<sup>fx</sup> -BG). For easy in vitro selection of human and bacterial cells that are transduced with the hybrid Ad vector, a bacterial sequence for the origin of replication can be added to the hybrid Ad vectors. An example of this is  
25  $\Delta$ Ad.AAV<sup>fx</sup> -Snori, in which a SNori sequence is added into the transgene cassette. This site allows for G418 selection on cells infected with  $\Delta$ Ad.AAV<sup>fx</sup> -SNori. This in vitro selection provides a tool to analyze the site of transgene integration and the flanking chromosomal regions. Fluorescent in situ hybridization (FISH) is an alternative method to confirm vector integration.

30

An advantage of the  $\Delta$ Ad chimeric vector for gene transfer is the efficient and stable integration of a large transgene cassette up to e.g. 22kb which is significantly larger than the capacity of retroviral vectors. This is of particular interest for gene therapy. For example, to ameliorate sickle cell anemia  $\Delta$ Ad.AAV<sup>fx</sup>  $\gamma$ globin, an expression transgene cassette with the gamma-globin gene that targets and integrates, can be inserted into bone marrow stem cells for long term expression of the gamma-globin gene.

To achieve site-specific gene integration, rep78 protein is used for transgene integration into the AAVS1 site (described in Example 1 D). However, this may silence transgene expression. To prevent the integrated transgene from being silenced by host genomic elements (such as positional effects or downstream enhancers), LCRs or insulator elements are incorporated into the transgene cassette.

The following examples are presented to illustrate the present invention and to assist one of ordinary skill in making and using the same. The examples are not intended in any way to otherwise limit the scope of the invention.

### **EXAMPLE I**

#### **20 NOVEL ADENOVIRAL VECTOR Ad.AAV**

##### **A. Integrating $\Delta$ Ad.AAV hybrid vectors devoid of all adenoviral genes.**

*In vitro* and *in vivo* studies with rAAV indicate that the only requirement for rAAV integration are the AAV ITRs and as yet unknown host cellular factors. It is thought that specific sequences or secondary structures present in AAV ITRs are prone to integration into host chromosomal DNA. In order to combine advantages of adenoviral vectors (high titer, high infectivity, large capacity) and the integration capability of AAV ITRs, AAV vector DNA with AAV ITRs flanking cassettes a secreted human placental alkaline phosphatase (SEAP) - neomycin phosphotransferase (neo) reporter gene cassette

(Alexander, I.E., et al. 1996, *Gene Therapy*, 7, 841-850) is incorporated into the E1-region of E1/E3 deleted adenoviral vectors (Ad.AAV1) (Figure 1, top).

## METHODS

5

### Production/Characterization of Viral Vectors

#### Plasmids:

- 10 The AAV1 vector cassette containing AAV ITRs and SEAP/neo expression units is obtained by AseI/ScaI digestion of the plasmid pALSAPSN (Alexander, I. E. et al, 1996, *Human Gene Therapy*, 7:841-50). The 4.4kb AAV vector fragment was cloned via NotI adapter linkers into pXJCL1 (Mirobix, Toronto, Canada) (pAd.AAV1). Another shuttle vector (pAd.AAV1-Δ2ITRs) lacking the AAV ITRs is generated by inserting the 3.7kb
- 15 AflII/BsmI fragment of pALAPSN into pXJCL1. For pAd.AAV1Δ1ITR, a construct is used where a spontaneous deletion in the left AAV ITRs between the A and A' regions has occurred. To create a second hybrid vector (Ad.AAV2), the AAVSNori cassette developed by E. Rutledge is used. AAV vector DNA obtained is from pASNori (Rutledge, E. A., Russell, D.W. 1997. *Journal of Virology* 71:8429-8436) as a 3.4kb
- 20 BsaI/ScaI fragment and inserted into the EcoRV site of pXCJL1. As it is generally known for AAV vector plasmids, the AAV ITRs are prone to rearrangements. To minimize deletions in these functional critical regions, all constructs for generation of hybrid vectors are assembled in low copy-number plasmids which are grown in *E. coli* Top10, JC811, or XLI Bluecells (Stragene, La Jolla, Calif.). Furthermore, after each
- 25 cloning step or large-scale plasmid amplification, both AAV ITRs are carefully mapped by restriction analysis with enzymes that cut inside or adjacent to the ITRs (BssHII, AhdI, SmaI, BglI, BsmI, AflII, and ScaI).

Adenoviruses:

First-generation viruses with the different transgene cassettes incorporated into the E1 region are generated by recombination of the pΔE1aSpl<sup>a</sup>- or pXCJL1-derived shuttle  
5 plasmids and pJM17 (Microbix) in 293 cells as described earlier (Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791). For each virus, at least 20 plaques are picked, amplified, and analyzed by restriction digest. Viruses containing two AAV ITDRs tend to rearrange within the ITRs, with other adenoviral sequences, or with adenoviral sequences present in the 293 cell genome. Only plaques from viruses with intact ITRs are amplified, CsCl  
10 banded, and titered as described earlier (Kay, M. A., et al. 1995. *Hepatology* 21:815-819; Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960). All virus preparations tested are negative for RCA and bacterial endotoxin (Lieber, A., et al. 1997. *Journal of Virology* 71:8798-8807). Virus is stored at -80°C in 10mM Tris-Cl, pH 7.5-1 mM MgCl<sub>2</sub>-10% glycerol.

15

To generate ΔAd.AAV, 293 cells are infected with Ad.AAV1 at an multiplicity of infection (MOI) of 25 and harvested 40h after infection. Cells are lysed in PBS by 4 cycles of freeze/thawing. Lysates are centrifuged to remove cell debris and digested for 30 min at 37°C with 500 units/ml DNaseI and 200 μg/ml RNaseA in the presence of  
20 10mM MgCl<sub>2</sub>. 5ml of lysate is layered on a CsCl step gradient (0.5ml - 1.5g/cm<sup>3</sup>, 2.5ml - 1.35g/cm<sup>3</sup>, 4ml - 1.25g/cm<sup>3</sup>) and ultracentrifuged for 2h at 35,000 rpm (rotor SW41). CsCl fractions are collected by puncturing the tube and are analyzed for viral DNA (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960; Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) or subjected to ultracentrifugation at 35,000 rpm for 18 hours  
25 in an equilibrium gradient with 1.32 g/cm<sup>3</sup> CsCl. The band containing the deleted viruses ΔAd.AAV is clearly separated (0.5cm distance) from other banded viral particles containing full-length ad.AAV genomes. ΔAd.AAV1 fractions are dialyzed against 10mM Tris-Cl, pH 7.5-1mM MgCl<sub>2</sub>-10% glycerol and stored at -80°C. The genome titer of ΔAd.AAV1 preparations is determined based on quantitative Southern analysis of viral  
30 DNA purified from viral particles in comparison to different concentrations of a 4.4kb AseI/ScaI fragment of pALSAPSN according to a protocol described earlier (Lieber, A.,

et al., 1996, *J. of Virology*, 70, 8782-8791). In total, the production of  $1 \times 10^{13}$  genome-particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of actual work.

5 Titers routinely obtained are in the range of  $3-8 \times 10^{12}$  genomes per ml. Assuming one genome is packaged per capsid, the genome titer equals the particle titer. The level of contaminating Ad.AAV1 is less than 0.1% as determined by Southern analysis, which is consistent with results obtained by plaque assay on 293 cells (fewer than 5 plaques per  $10^6$  total genomes). The primers used for sequencing the left and right ITR-vector-junction are

10

5'GGCGTTACTTAAGCTAGAGCTTATCTG, and  
5'CTCTCTAGTTCTAGCCTCGATCTCAC.

15 The recombinant AAV virus stock containing the SEAP/neo cassette (AV2/ALSAPSN, [Alexander, I. E. et al, 1996, *Human Gene Therapy*, 7:841-50] used in these studies were obtained from Dusty Miller (FHCRC, Seattle). The stock was free of contaminating replication competent AAV ( $<50$  particles/ml) and wildtype adenovirus ( $<100$  particles/ml). The genome titer of the virus stock was obtained by quantitative Southern Blots as described by Russell et al. (Russell, D. et al. 1994 *Proc. Natl.Acad.Sci. USA*  
20 91:8915-8919).

#### Electron Microscopy:

25 For examination of viral particles in the transmission electron microscopy studies, CsCl-purified virions are fixed with glutaraldehyde and stained with uranyl acetate as described previously (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960).

#### **RESULTS**

30 During replication of these hybrid vectors in 293 cells, a 5.5kb genome ( $\Delta$ Ad.AAV1) is efficiently generated and packaged into adenovirus (Ad5) capsids. The  $\Delta$ Ad.AAV1

genome contains the left adenovirus ITR and the packaging signal followed by the AAV-vector cassette and a duplicate of the adenoviral packaging signal and ITR in reverse orientation (Figure 1, bottom). The hybrid vector is devoid of all viral genes, thus eliminating toxic effects and the elicitation of cellular immune responses. The spontaneous formation of the small hybrid vector genome  $\Delta$ Ad.AAV1 requires the presence of two intact AAV ITRs and does not occur with partly deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette.

Hybrid vectors containing different transgenes:

To construct a hybrid vector with a transgene that can be detected in situ the SEAP/neo expression unit in Ad.AAV1 is replaced by the *E. coli*  $\beta$ -galactosidase gene. This hybrid vector is named  $\Delta$ Ad.AAV1. During generation of the corresponding plasmid constructs the AAV ITR sequences tend to rearrange and abolish their functional properties. This problem can be circumvented by using low copy number plasmids as cloning vectors grown in bacteria strains depleted for all recombination proteins (e.g.JC811). Furthermore, the intactness of both AAV ITRs after each cloning step can be examined for characteristic endonuclease digestion. Recently, another hybrid vector  $\Delta$ Ad.AAVINori has been generated which contains the neo gene under the control of both the simian virus 40 (SV40) early promoter and the transposon 5 (Tn5) promoter for expression in human and bacterial cells, as well as the p15A bacterial replication origin with the direction of the leading strand DNA synthesis opposite that of neo gene transcription. Thus, SNori can be used for G418 selection of integrated vector in eukaryotic cells as well as for rescue of vector together with flanking host DNA after integration. The recovered plasmids can be propagated in *E. coli* under selection with kanamycin due to the bacterial origin and the neo gene. SNori containing vectors allow a rapid estimation of total integration events based on the number of G418 resistant colonies. Moreover, vector DNA together with flanking chromosomal DNA can be rescued as plasmids from single G418 resistant clones and can be used for sequencing to determine integration junctions. Both hybrid vectors are produced at a titer of about

$3 \times 10^{12}$  genomes per ml. The ratio of genome titer to transducing particles for  $\Delta$ Ad.AAVBG is  $\sim 200:1$  based on  $\beta$ -Gal expression.

## DISCUSSION

5  $\Delta$ Ad.AAV1 could spontaneously form during adenovirus replication. Another possible mechanism of  $\Delta$ Ad.AAV1 formation is based on the unique mechanism by which adenovirus replicates its genome (van der Vliet, B., 1995, *In w. Doerfler, et al. (eds.) vol. 2 p. 1-31, Springer-Verlag, Berlin*) (see Figure 1). Ad DNA replication is initiated by the  
10 TP/pTP (terminal protein) that binds to specific sites within the ITRs on both ends of the linear genome and catalyzes, in complex with Ad pol, the binding of the 5' CTP, the first nucleotide of the daughter strand. DNA synthesis proceeds in a continuous fashion to the other end of the genome (Figure 1A). Only one of the DNA strands serves as template. One of the replication products is a single-stranded DNA that circularizes through  
15 annealing of its self-complementary ITRs. The resulting duplex "panhandle" has the same structure as the termini of the duplex viral genome that allows the binding of pTP and the initiation for synthesis of a complementary strand using the single-stranded "panhandle" molecule as template (Figure 1C). In the case of Ad.AAV1, the Ad pol synthesizes the single strand of the adenoviral genome starting from the left Ad ITR until  
20 it reaches the second AAV ITR. During synthesis of the second AAV ITR a certain percentage of the single-stranded molecules form a loop hybridizing to the complementary region within the first AAV ITR that was replicated earlier, allowing Ad pol to use the same viral DNA strand to read back towards the left ITR (Figure 1B). The resulting "panhandle" structure can be resolved in a similar way as a full-length  
25 intermediate shown in Figure 1C, generating a double stranded, linear molecule with the above described structure that can be packaged into Ad virions. The ratio of viral DNA to protein concentration in purified  $\Delta$ Ad.AAV1 particles is comparable to that obtained from Ad.AAV1 particles. This indicates that despite the smaller size, only one  $\Delta$ Ad.AAV1 genome is packaged, resulting in particles with a lighter buoyant density  
30 ( $\sim 1.32 \text{ g/cm}^3$ ). Electron microscopy demonstrates the icosahedral shape of  $\Delta$ Ad.AAV1 particles (Figure 2). Staining with uranyl acetate causes the central viral cores to appear



electron dense.  $\Delta$ Ad.AAV1 virions have only a spotted luminal dark staining as expected with only one 5.5kb genome being packaged per capsid.

**B. In vitro  $\Delta$ Ad.AAV1 production:**

**Characteristics of deleted adeno-AAV vectors ( $\Delta$ Ad.AAV):**

A number of experiments to clarify the mechanisms of  $\Delta$ Ad.AAV genome formation are carried out. Specifically, the presence of two intact AAV ITRs flanking a reporter gene cassette is required for the effective formation of  $\Delta$ Ad.AAV genomes. This process does not occur with partially deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette. Furthermore, in vitro transduction studies are performed with different genome titers of  $\Delta$ Ad.AAV1, Ad.AAV1, and Ad.AAV1- $\Delta$ 2ITRs, (lacking the two AAV ITRs) which determine the number of G418 resistant colonies that formed after 4 weeks of selection (Table I).

$\Delta$ Ad.AAV1 is routinely produced at a high titer ( $5 \times 10^{12}$  genomes per ml with  $>10^4$  produced genomes per 293 cell) and at a high purity with less than 0.1% contaminating full length Ad.AAV1 genomes by a technique normally used for amplification and purification of recombinant adenovirus.

**In vitro transduction studies with hybrid vectors on CD34+ cells and erythroleukemia cells:**

In order to test whether the hybrid vectors allow for gene transfer into cell types, that have to be targeted for sickle cell therapy, infection/transduction studies are performed using CD34+ enriched human bone marrow cells, derived from mobilized peripheral blood and the human erythroleukemia cell line K562 which express  $\epsilon$  and  $\gamma$  globin genes.

**METHODS**

Cell Culture:

SKHep1 cells (HTB-52, American Type Culture Collection, Rockville, MD), an endothelial cell line derived from human liver [Heffelfinger, S.C., et al., 1992, *In vitro* 5 *Cell Dev. Biol.* 28A, 136-4-142], are grown in high-glucose Dulbecco's modified Eagle medium with 10% fetal calf serum. SKHep1 cells are analyzed for integrated AAV provirus by Southern analysis of genomic DNA using the AAV1 wild type genome obtained from pAAV/Ad (Samulski, R. J., et al. 1989. *Journal of Virology* 63:3822-3928) (gift from David Russell, University of Washington) as a probe. No specific bands are 10 detected in undigested genomic SKHep1 DNA or after digestion with HindIII. For viral infection, confluent cells are incubated with different viral doses for 2 hours, followed by intensive washing. For G418 selection, 24h after infection with  $\Delta$ Ad.AAV1, SKHep1 cells are trypsinized and plated at different dilutions under G418 selection (900 $\mu$ g/ml active compound, Boehringer-Mannheim, Germany). G418 containing culture medium is 15 changed every 3 days. The number of colonies with  $>10^4$  cells is counted after 4 weeks of selection and divided by the number of initially seeded cells. This ratio is used to express the integration frequency of  $\Delta$ Ad.AAV1. Single colonies are obtained by limiting dilutions of infected cells in 96 well plates. Colonies are expanded to  $1 \times 10^6$  cells in the presence of G418. Immunofluorescence analysis for adenoviral proteins expressed in 20 SKHep1 cells 3 days post-infection is performed as described earlier [Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791].

RESULTS

25 293 cells are infected with the first generation vector Ad.AAV1. During replication of Ad.AAV1, the small  $\Delta$ Ad.AAV1 genome forms spontaneously and is packaged into adenovirus capsid. At 36 hours after infection cells are harvested and virus is released by several cycles of freeze/thawing. The mixture of Ad.AAV1 and  $\Delta$ Ad.AAV1 particles in the cell lysate is then separated by ultracentrifugation in a CsCl step gradient. Due to its 30 lighter buoyant density, the band containing the  $\Delta$ Ad.AAV1 particles is clearly separated (0.8cm distance) from the band containing full-length virus (Lieber, A., et al. 1999. *J*

*Virology* 73:9314-24).  $\Delta$ Ad.AAV1 is purified further by an additional CsCl equilibrium gradient and is stored in 10mM Tris pH7.5, 10% glycerol, 1mM MgCl<sub>2</sub> in 80°C. In total, the production of  $2 \times 10^{13}$  (genome) particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of work. All functions for  $\Delta$ Ad.AAV1 replication and particle formation are provided from

5 Ad.AAV1 genomes amplified in the same cell. The efficiency of vector production measured on a genome-per-cell-basis is comparable or higher than labor-intensive, newer techniques for rAAV production, which have not yet been proven to be reliable. The estimated ratio of transducing/genome titer for  $\Delta$ Ad.AAV1 is 1:200 (based on SEAP expression at day 3 post-infection), whereas for the average rAAV preparation, it is in the

10 range of  $1:10^3$  to  $1:10^4$ .  $1 \times 10^5$  confluent SKHepl cells are infected with different MOIs of rAAV1 (stock:  $1 \times 10^{10}$  genomes per ml),  $\Delta$ Ad.AAV1 (stock:  $5 \times 10^{12}$  genomes per ml), Ad.ADAV1 (stock:  $1 \times 10^{13}$  genomes per ml), and AdAAV1.2ITR (stock:  $9 \times 10^{12}$  genomes per ml), in a volume of 100ml 24 hours after infection, cells are washed, trypsinized, and plated at different dilutions. G418 is added 24 hours after plating and selection is

15 performed for 4 weeks. G418 resistant colonies contain on average  $>5 \times 10^4$  cells (at least 16 cell divisions). A significant number of small colonies visible at 2 weeks post-infection do not survive continued selection, probably due to episomal vector expression. Cells infected with first-generation adenoviruses with MOIs greater than  $1 \times 10^4$  develop CPE during the first week of selection. The rAAV titer is not high enough to perform

20 infection studies with MOIs greater than  $10^4$ . The colony formation is expressed as percentage of the number of colonies after selection to the number of cells initially seeded for selection (Table I).

**TABLE I**

25

**Formation of G418 resistant colonies after infection with hybrid viruses in comparison with rAAV.**

| MOI<br>(genomes per<br>cell) | Formation of G418 resistant colonies in % (SEM)<br>(after 4 weeks of selection) |
|------------------------------|---|
|------------------------------|---|

|        | rAAV1      | .Ad.AAV1   | Ad.AAV1    | Ad.AAV1<br>.2ITRs |
|--------|------------|------------|------------|-------------------|
| $10^1$ | 0          | 0          | 0          | 0                 |
| $10^2$ | 0          | 0          | 0          | 0                 |
| $10^3$ | 2.7 (1.6)  | 1.3 (1)    | 5.4 (3.0)  | 0                 |
| $10^4$ | 90.8 (7.0) | 48.0 (8.9) | 12.9 (7.2) | 0                 |
| $10^5$ | N/A        | 93.1 (5.4) | 3.8 (2.1)  | 0                 |
| $10^6$ | N/A        | 100        | 0          | 0                 |
| $10^7$ | N/A        | 100        | 0          | 0                 |

N = 3 (SEM is indicated in parentheses.)

K562 cells are infected with different MOIs of  $\Delta$ Ad.AAVBG ( $1-10^8$  genomes per cell):

- 5 Three days after infection, the total number of viable cells (based on Trypan blue staining) and the percentage of infected cells (based on X-Gal staining) are determined for all MOIs. The results are presented in Figure 4A.

Initial integration studies:

10

K562 cells are incubated with  $\Delta$ Ad.AAVSNori at an MOI of  $2 \times 10^5$  genomes per cell and the colonies that formed after 4 weeks of G418 selection are counted in 96 well plates. G418 resistant colonies contain on average  $>5 \times 10^4$  cells which means that the original cell underwent at least 16 cell divisions.

15

Infection studies with .Ad.AAVBG ( $1-10^8$  genomes per cell) on CD34+ cells:

Cell infection on CD34+ are as described for K562 cells. CD34+ cells are cultured in IMDM supplemented with 20% FCS, kit ligand (stem cell factor-SCF) (100ng/ml), and  
 20 IL-3 (100ng/ml). Since a number of reports suggest that specific cytokines like GM-CSF or M-CSF which induce stem cell differentiation can stimulate integrin expression and may therefore affect internalization of Ad5 vectors, infection rates are compared with

Ad5 based hybrid vectors on CD34+ cells cultured with and without pre-stimulation with GM-CSF (50ng/ml) or M-CSF (50U/ml). The number of infected cells is counted based on X-Gal staining at day 3 after infection. To test for dose-dependent toxicity, viable cells are counted (based on trypan blue exclusion) at day 3 post-infection. Furthermore, whether high viral doses affect the ability of CD34+ cells to differentiate in methyl cellulose colony assays in presence of IL-3 and SCF is analyzed. The results are expressed as viable cells/X-Gal positive cells vs MOI (see Figure 4B).

## DISCUSSION

10

The above data demonstrate that  $\Delta$ Ad.AAV transduces stably an immortalized human cell line with a low frequency comparable to rAAV, however, transduction rates could be scaled up to 100% by using greater MOIs of  $\Delta$ Ad.AAV1, which is produced at higher titers than rAAV1. In contrast to infection with the first-generation vector, Ad.AAV1, infection with  $\Delta$ Ad.AAV1 is not associated with dose-dependent cytotoxicity because no viral proteins are expressed from these vectors in transduced cells. Furthermore, viral proteins present in the incoming  $\Delta$ Ad.AAV1 particles are not problematic in the dose range use. The comparison of transduction rates of  $\Delta$ Ad.AAV/Ad.AAV1 with the vector lacking AAV ITRs, Ad.AAV- $\Delta$ 2ITRs, supports the hypothesis that the presence of two intact AAV ITRs is crucial for hybrid vector integration.

The data demonstrate that the leukemia cell line can be infected at ~90% efficiency with Ad5 based hybrid vectors at MOIs  $2 \times 10^5$  genomes per cell without significant toxic side effects. However, this dose is still ~100 times greater than the dose necessary to infect 100% of HeLa cell, hepatoma cells, primary hepatocytes and other cell lines generally considered as permissive for Ad5 vector infection.

Since viral DNA in cells infected with  $2 \times 10^4$  genomes (or 100 transducing particles per cell) should be lost after 7 cell divisions, the presence of G418 resistant cells in the observed colonies suggests that  $\Delta$ Ad.AAVSNori genomes are integrated into or stably associated with the host genome. Based on the number of G418 resistant colonies one

out of 25,000  $\Delta$ Ad.AAVSNori genomes integrates stably into K562 cells. This is in agreement with the results obtained earlier with  $\Delta$ Ad.AAV1 in SK Hep1 cells.

5 The maximal dose used for infection of CD34+ cells ( $1 \times 10^8$ ) results in X-Gal staining of only ~10% of cells independently of GM-CSF/M-CSF. This demonstrates the obvious inability of Ad5 to infect CD34+ cells and is probably caused by the absence of specific receptors and/or integrins on the cell surface. CD34+ cells tolerate a large range of viral doses ( $1-10^7$ ) without obvious effects on cell viability and total cell number. This is not surprising because in order to develop toxic side effects adenovirus has to enter the cell and express viral genes. Hybrid vectors can be produced at titers of  $5 \times 10^{12}$  genomes per ml. Thus, the maximal MOI that can be used for infection (of  $10^4$  cells) is  $\sim 5 \times 10^8$  (in 100 $\mu$ l storage buffer). Based on the infection studies with  $\Delta$ Ad.AAVBG this dose may not be sufficient to efficiently transduce CD34+ cells and to obtain an appreciable number of G418 resistant colonies.

15

### C. In Vivo Properties Of $\Delta$ Ad.AAV1:

Viral DNA is labeled with BrdU during virus amplification to investigate cellular/nuclear vector uptake in situ. For transduction studies, confluent SKHep1 cells (a human endothelial cell line) are infected with 2000 genomes  $\Delta$ Ad.AAV1 or Ad.AAV1 per cell. BrdU tagged viral DNA is detected in 100% of nuclei at 3 hours post-infection for both viruses indicating efficient cellular and nuclear uptake of hybrid virus DNA.

### RESULTS

25

The  $\Delta$ Ad.AAV1 vector transduces a cell in vitro forming G418 resistant colonies with an efficiency of 17 or 58%, after infection with an MOI of  $1 \times 10^3$  or  $1 \times 10^4$  genomes per cell, respectively. Approximately  $2 \times 10^4$   $\Delta$ Ad.AAV1 genomes are required to yield one stable transfectant. Since all stable colonies contain integrated  $\Delta$ Ad.AAV1 vector DNA, this number reflects the minimal integration frequency of  $\Delta$ Ad.AAV1 in SKHep1 cells which is comparable with that from rAAV (Rutledge, E. A. et al., 1997, *Journal of Virology*,

30

71:8429-36). The number of G418 resistant colonies does not necessarily represent the total frequency of integration events because not all integrated copies express neomycin phosphotransferase, due to chromosomal position effects or incomplete integration.

- 5 The absence of adenoviral gene products in  $\Delta$ Ad.AAV1 transduced cells at day 3 post-infection is demonstrated by immunofluorescence with antibodies to the major late proteins (hexon, fiber) and early proteins (DBP.E4-orf6). Expressed adenoviral proteins are detected only in cells infected with Ad.AAV1. The fact that cells infected with  $\Delta$ Ad.AAV1 do not express potentially cytotoxic adenoviral proteins is important.

10

- While an MOI of  $1 \times 10^4$  genomes per cell of the first generation vector Ad.AAV1 induce cytopathic effects in SKHep1 cells at day 3 p.i., no toxic side effects are observed when SKHep1 cells are infected with  $\Delta$ Ad.AAV1 at a dose of up to  $1 \times 10^8$  genomes per cell. Since the transduction efficiency is clearly dose dependent,  $\Delta$ Ad.AAV1 (which can be  
15 produced at titers of  $>5 \times 10^{12}$  genomes/ml) is able to stably transduce Ad5 permissive cell lines or tissues with a 100% efficiency without associated toxicity.

- Southern analysis indicates that  $\Delta$ Ad.AAV1 integrates randomly as head-to-tail tandem repeats into the host cell genome via the right AAV ITR, whereas the other junction with  
20 the chromosomal DNA is variable and occurs somewhere within the transgene cassette. In order to confirm the integrated status of  $\Delta$ Ad.AAV1 DNA, high-molecular-weight chromosomal DNA is separated by pulse field gel electrophoresis (PFGE), followed by Southern analysis with a SEAP specific probe (Figure 3). Undigested DNA from control SKHep1 cells give an endogenous SEAP signal that co-migrates with chromosomal DNA  
25 just below the well (lanes 1 and 5). No high-molecular weight episomal forms of  $\Delta$ Ad.AAV1 DNA are detected, whereas a distinct 35 kb band is visible in DNA from SKHep1 cells isolated 3 days after infection with first generation adenovirus, Ad.AAV1 (lanes 4 and 13). Digestion with EcoRI reveals the 4.4kb fragment, which is specific for integrated tandem copies of the AAV cassette (lanes 8 and 12). To eliminate the  
30 possibility that chromosomal DNA is trapped in the well, DNA samples are digested with intron-encoded endonucleases PI-SceI or I-CeuI (Gibco-BRL, Grand Island, NY) with a

sequence specificity or more than 11bp or 9bp respectively. Digestion with PI-SceI yields a >2mb endogenous SEAP signal in SKHep1 cells (lane 2) and an additional signal in the range of ~1mb in G418 resistant colonies transduced with  $\Delta$ Ad.AAV1 (lane 7). I-CeuI digestion results in a smear between 250-1000 kb in  $\Delta$ Ad.AAV1 transduced SKHep1-cells (lanes 10, 11) indicating random integration, whereas a high-molecular weight band specific for the endogenous SEAP gene is observed in control SKHep1 cells (lane 9).

One day after intraportal infusion of  $1 \times 10^{12}$   $\Delta$ Ad.AAV1 genomes in C57Bl/6 mice, BrdU labeled vector genomes can be detected in 85% hepatocytes (Lieber, A., et al. 1999. *J Virol* 73:9314-24). Hepatocellular DNA analysis performed at 2 months post-infusion reveals  $\Delta$ Ad.AAV1 DNA integrated with an average of 0.5 copies per cell into the mouse genome (Lieber, A., et al. 1999. *J Virol* 73:9314-24). To assess potential side effects of intraportal  $\Delta$ Ad.AAV1 infusion, serum glutamic pyruvic transaminase (SGPT), a sensitive marker for hepatocellular injury, is measured for 7 consecutive days post-infusion in combination with histological analysis of liver sections. No significant elevation in SGPT levels, or histological abnormalities are detected after intraportal infusion of  $1 \times 10^{12}$  or  $1 \times 10^{13}$   $\Delta$ Ad.AAV1 genomes, whereas infusion of the same dose of full-length Ad.AAV1 vector is associated with severe hepatotoxicity or fatal outcome. This suggests that the dose of  $\Delta$ Ad.AAV1 administered to mice can be increased to obtain higher transduction efficiencies in vivo without adverse side effects, which is not possible for first generation adenoviruses. Importantly,  $\Delta$ Ad.AAV1 transduced quiescent hepatocytes in vivo, which suggests that integration of hybrid vector DNA may not require cell proliferation. Recently, more detailed in vivo transduction studies with Ad.AAV1 and  $\Delta$ Ad.AAV1 have been performed in Balb/c mice to study whether the absence of adenoviral gene expression in cells infected with  $\Delta$ Ad.AAV1 can avoid an anti-viral immune response and can prolong vector persistence. In this mouse strain, vector DNA is cleared from the liver at 4-6 weeks after infusion with first generation adenoviruses, mostly due to a CTL response against viral proteins produced in transduced cells. Vector DNA is analyzed by genomic Southern Blot of hepatic DNA at 12 weeks



after infusion of  $1 \times 10^{12}$  genomes Ad.AAV1 or  $\Delta$ Ad.AAV1. At this time point, no vector specific signal is detectable in hepatic DNA from mice infused with the first generation vector Ad.AAV1, while  $\sim 0.3$  copies of  $\Delta$ Ad.AAV1 genomes per cell are present in livers of mice that received the hybrid vector, again indicating the superior in vivo properties of the hybrid vector.

#### D. Effects of Rep Coexpression on $\Delta$ Ad.AAV Integration

##### Rep expression after plasmid transfection:

In order to test whether Rep expression enhances site-specific integration of  $\Delta$ Ad.AAV1 in human cells, a series of Rep expression plasmids are constructed.

#### METHODS

The Rep ORF 68/78 (nt 285-2313) including the internal p19 and p40 promoters is obtained from pAAV/Ad (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) by digestion with BsaI/BsrI. This fragment deleted for the AAV p5 promoter is cloned via adapter linkers under RSV or PGK promoter in front of the bovine growth hormone polyadenylation signal (bPA) into pAd.RSV or pAd.PGK (Lieber, A., and Kay, M.A., 1996, *J. of Virology*, 70, 3153-3158; Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) correspondingly.

#### RESULTS

The resulting plasmids (pRSVrep, pPGKrep) are transfected into 293 cells or SKHep1 cells, most of the Rep proteins expressed from the heterologous promoters (RSV or PGK) are Rep 68 and Rep 78, while transfection of the rep gene under aP5 promoter (pAAV/Ad) results in predominant Rep 52/40 expression. Thus, transfection of pRSVrep and pPGKrep is more pronounced suggesting a strong transactivation of AAV promoters

by Ela which is produced in 293 cells. This result indicates that minimum expression of rep proteins is necessary to avoid interference with adenovirus replication.

Rep-mediated site-specific integration of  $\Delta$ Ad.AAV1.

5

The potential for site-specific integration is an important characteristic of the novel Ad.AAV vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration, which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

A preliminary test can be performed to confirm the functional activity of Rep 68/78 expressed from pRSVrep to mediate site-specific integration of  $\Delta$ Ad.AAV1 (Figure 5 and 6). Human SKHep1 cells are transfected with pRSVrep or control plasmid (pRSVbGal (Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) (transfection efficiency was ~20%), followed by infection with  $\Delta$ Ad.AAV (2000 genomes per cell). Three days after infection, cells are trypsinized, embedded in agarose, lysed in situ, digested with I-CeuI (an intron-encoded endonuclease with a recognition sequence of more than 10nt), subjected to pulse field gel electrophoresis in 1% agarose gel, and analyzed by Southern Blot. Hybridization with a probe covering the AAVS1 integration site (1.7kb EcoRI/BamHI fragment from the chromosome 19 locus (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) ) reveals an AAVS1-specific band (~240kb) in I-CeuI digested DNA from cells after control plasmid transfection (pCo) +  $\Delta$ Ad.AAV1 infection. An additional signal in the range of 280kb appears in rep expressing cells infected with  $\Delta$ Ad.AAV1

(pRSVrep + VAd.AAV1) indicating a site-specific insertion into the AAVS1 site in a certain percentage of cells. The presence of vector DNA in this 280kb band is confirmed by rehybridization of the same filter with a transgene (SEAP) specific probe. Randomly integrated  $\Delta$ Ad.AAV1 vector appears as a diffuse SEAP signal in the range 280-680kb  
5 (pCo+ $\Delta$ Ad.AAV1, pRSVrep+ $\Delta$ Ad.AAV1). The specific ~1.9 mb band on blots hybridized with the SEAP probe represents an I-CeuI fragment containing the endogenous human SEAP gene.

Incorporation Rep 68/78 function into hybrid vectors to stimulate site-specific integration  
10 Rep overexpression inhibits adenovirus DNA replication, prohibiting the generation of rep expressing Ad vectors using conventional strategies. To solve this problem, significant Rep 68/78 expression from the hybrid vector in virus producer (293) cells must be prevented while maintaining transient Rep expression in target cells (HSC) to mediate site-specific integration. Our hypothesis is that the specific structure of the  
15  $\Delta$ Ad.AAV hybrid virus can be used to bring the rep gene 68/78 into a transcriptionally active position under control of a HSC specific promoter only at late stages of virus replication in 293 cells. This will allow amplification of the hybrid vector in 293 cells, generating high titer virus which activates the incorporated Rep 68/78 functions only in HSC. The general outline of our strategy to produce Rep expressing hybrid vectors is  
20 illustrated in Figure 7. The rep/transgene cassette is assembled based on the left-hand shuttle plasmid used for recombinant adenovirus production. The gene encoding Rep 68/78 is cloned in 3' to 5' orientation in front of a transgene expression cassette flanked by AAV ITRs. Between the transgene cassette and the right AAV ITR an HSC-specific promoter is inserted with direction towards the adenoviral E2, E3, and E4 genes. The  
25 recombinant genome is produced by recombination in E.coli and transfection into 293 cells generates virus (Ad.AAV-rep). The specific structure of  $\Delta$ Ad.AAV with duplicated sequences flanking the AV ITRs is used to bring the rep gene into a transcriptionally active position under control of a HSC specific promoter only during late stages of viral DNA replication in 293 cells. During amplification of Ad.AAV-rep, the smaller genome  
30  $\Delta$ Ad.AAV-rep is formed and packaged into particles, which can be separated by ultracentrifugation in CsCl gradients. The specific structure of  $\Delta$ Ad.AAV-rep brings the

rep gene into 5'↔3' orientation in relation to the HSC specific promoter, allowing rep transcription in target cells. After transduction of HSC with purified ΔAd.AAV-rep particles, rep expression is activated and mediates rescue of the AAV-ITR/transgene cassette from the adenoviral vector backbone and site-specific integration. The hypothesis is that Rep-mediated integration into AAVS1 occurs via the right or both AAV/ITRs causing the rep gene to become separated from the hepatocyte-specific promoter once the vector is integrated (Figure 7). Therefore, rep expression should be only transient without critical cytotoxic side effects on the host cell.

10 Promoters that can regulate rep expression:

Potential candidate promoters to drive rep expression with high specificity for HSC and minimal activity in 293 cells are the 454nt CD34 promoter (Krause, D.S., et al., 1997, *Experimental Hematology*, 25, 1051-1061; Yamaguchia, Y. et al., 1997, *Biochimica et*  
15 *Biophysica Acta.*, 1350:141-6), the 300nt HS 40 enhancer (Chen, H.L., et al., 1997, *Nucleic Acids Res.* 25, 2917-2922) or a 3kb CD34 enhancer (May, G. et al., 1995, *EMBO J.*, 14:564-74) in combination with an initiator, or the HIV LTR. An optimal promoter is selected based on studies of transient reporter gene expression after plasmid transfection in 293 cells and hepatocytes. All promoters to be tested are cloned in front of the human  
20 α<sub>1</sub>-antitrypsin (hAAT)- bovine growth hormone polyadenylation signal (bPA) into the adenoviral shuttle plasmid pCD2 (pAd.-hAAT). Promoter activity can be tested in transient plasmid transfection assays in CD34+ and 293 cells. The promoter with the highest hAAT levels in CD34+ or K562 cells and the lowest hAAT expression in 293 cells is selected for further studies. If high background expression in 293 cells from these  
25 promoters is seen, insulators to shield HSC-specific promoters from the Ela enhancer which is still present in Ad shuttle plasmids can be utilized.

Rep genes:

30 The large Rep 68/78 proteins are sufficient to mediate rescue and site-specific integration. Unregulated Rep 52 and Rep 40 expression from the AAV p19 promoter

located within the ORF of Rep 68 and 78 must be prevented because production of these smaller Rep proteins in 293 cells will affect cell viability and adenoviral DNA synthesis. To do this, constructs obtained from Surosky et al., containing a mutated Rep 52/40 start codon to express Rep 68 and 78 individually under CMV promoter can be used. The 293  
5 cells transiently expressing Rep68 or Rep 78 from these constructs can be coinfectd with  $\Delta$ Ad.AAV1 (infection 24 hours after pCMVRep transfection, MOI  $2 \times 10^5$  genomes/cell). Three days after  $\Delta$ Ad.AAV1 infection, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE as described earlier. Efficient Rep mediated excision of the AAV cassette and site-specific integration without flanking adenoviral  
10 sequences are expected and the plasmids pCMVRep68 or pCMVRep78 can be used as a source for the corresponding rep genes and clone them into hybrid vectors.

Vectors:

15 The rep/transgene cassette can be assembled based on pXCJL (Microbix, Toronto). A set of control hybrid vectors can be generated with the AAV-ITR-transgene cassette only without the rep gene. The recombinant Ad.AAV-rep genome can be generated by recombination of the left hand shuttle plasmids with pCD1, a pBHG10 (Microbix, Toronto) derivative, which contains the Ad5 genome deleted for the E1/E3 regions in  
20  $recA^+$  E. coli (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Compared to the standard technique based on plasmid recombination in 293 cells, this approach has the advantage that plaques with recombinant virus appear 3 times faster and the production of illegitimate recombinants is minimized. This allows efficient viral DNA amplification and packaging to occur before Rep expression reaches levels that are potentially  
25 inhibitory for adenoviral replication. The critical variables in maximizing the output of the vector deleted for all adenoviral viral genes are the initial multiplicity of infection and the time of harvesting. These parameters can be optimized for production of  $\Delta$ Ad.AAV-rep hybrid vectors. A number of  $\Delta$ Ad.AAV vectors can be constructed incorporating rep gene. Cryptic promoter and enhancer elements present in the 5'-342nt of the adenoviral  
30 genome can interfere with transgene expression from the heterologous promoters. This is crucial for the strategy to avoid rep expression from  $\Delta$ Ad.AAV-rep genomes in 293 cells.

To ensure efficient transgene expression, insulator fragments such as the chicken beta-globin insulator can be used with a selected promoter, constitutive or inducible.

Rep protein co-packaging:

5

As an alternative to producing hybrid vectors containing the rep 68/78 gene, studies are designed to see whether Rep protein can be co-packaged into  $\Delta$ Ad.AAV capsids and whether these co-packaged Rep molecules are sufficient to mediate rescue and site-specific integration of the AAV-ITR-transgene cassette. Our hypothesis is that the Rep  
10 68/78 binds to the Rep binding site (RBS) present in double-stranded  $\Delta$ Ad.AAV genome and that this complex is co-packaged into adenoviral capsids which are spacious enough to accommodate extra proteins. Based on protein/DNA ratio analysis performed previously in purified particles that only one 5.5kb  $\Delta$ Ad.AAV1 genome is packaged per capsid. This is confirmed by electron-microscopy of  $\Delta$ Ad.AAV1 particles, which reveals  
15 only spotted electron-dense staining associated with viral cores and extended free luminal space (see Figure 2).

293 cells are transfected with plasmids expressing Rep 68/78 under the CMV promoter and the kinetics of rep expression is determined by Western Blot with cell lysates  
20 collected at different time points after transfection. Next, these 293 cells are infected with Ad.AAV (MOI 1, 10, 100 pfu/cell) at specific time points after transfection of Rep plasmids depending on the Rep expression kinetics (e.g. 3, 6, 12, 24 . . . hours after transfection). It is important to time Ad.AAV infection exactly because viral DNA replication must be taking place or finished before Rep production reaches peak levels.  
25 In general, adenovirus DNA replication in 293 cells (infected with MOI 10) is maximal at 18 hours post-infection, followed by production of structural proteins, packaging of viral genomes, and breakdown of cellular membrane structures (which is concluded ~36-48h p.i.) (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; van der Vliet, B., 1995, *In* w. Doerfler, et al. (eds.) vol. 2  
30 p. 1-31, Springer-Verlag, Berlin). Viruses are collected 48h after infection and banded by CsCl ultracentrifugation. Viral material from purified bands corresponding to

ΔAd.AAV is lysed, DNase-treated (to liberate DNA associate Rep) and subjected to immunoprecipitation-Western Blot with Rep specific antibodies to detect co-packaged Rep. Based on theoretical calculations assuming that two Rep molecules bind per Ad genome, ~1-10ng Rep proteins is expected from Lysates of  $10^{10}$  particles, which is within  
5 the range of detectability by Western Blot. Alternatively, co-packaged Rep may be detected based on its functional activity to mediate rescue and site-specific integration of the AAVITR transgene cassette. To test whether functional Rep protein is co-packaged into hybrid vector particles, CsCl purified ΔAd.AAV1 particles generated in 293 cells co-expressing Rep after Ad/AAV1 infection (ΔAd.AAV1+Rep) can be used for transduction  
10 studies. Three days after ΔAd.AAV1+Rep infection of the human cell line K562, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE. If efficient Rep-mediated site-specific integration of excised AAV cassettes is successful, then other ΔAd.AAV+Rep hybrid vectors with β-Gal and SNori as transgenes can be produced.

15 Integration studies with Rep vectors in erythroid cells:

The hypotheses behind the rational of a rep-expressing hybrid vector (ΔAd.AAV-rep) are: (1) transient Rep co-expression from ΔAd.AAV-rep vectors can enhance site-specific vector integration in human cells and (2) integration occurs via the AAV ITR(s)  
20 without the rep gene, which is placed outside the AAV cassette, thus eliminating rep expression upon vector integration. To test hypothesis 1, transduction frequencies of ΔAd.AAV/rep versus ΔAd.AAV vectors can be compared based on the formation of G418 resistant colonies and quantify site-specific integration events at different time points after infection of human and mouse cells by PFGE and PCR. To test hypothesis 2,  
25 the structure of integrated vector in transduced cell populations and single clones can be delineated by Southern analysis and by sequencing of vector/chromosomal DNA junctions. These studies can be performed with ΔAd.AAV-rep, ΔAd.AAV, and ΔAd.AAV+Rep (copackaged protein) in human K562 or HEL (for AAVS1 integration) and mouse MEL cell lines.

30

Cells infected with  $\Delta$ Ad.AAV-SNori,  $\Delta$ Ad.AAV-SNori+Rep or  $\Delta$ Ad.AAV-Snori-rep can be subjected to G418 selection. The number of G418 resistant colonies determined after 4 weeks of selection in relation to the number of initially infected cells. The selection process for colonies that did not survive continued selection due to potential rep-mediated cytotoxicity or episomal vector expression can be monitored. If rep expression from  $\Delta$ Ad.AAV-SNori rep does not affect cell viability and proliferation, then more G418 resistant colonies should appear in  $\Delta$ Ad.AAV-SNori-rep and  $\Delta$ Ad.AAV-Snori+Rep. The structure of integrated vector can be determined by Southern Blot and sequencing of integration junctions.

To uncover a potential selection bias against rep producing cells after transduction with  $\Delta$ Ad.AAV/rep, site-specific and random vector integration events can be quantitated in cellular DNA isolated from cell populations at different time points after infection (e.g. 0.5, 1, 3, 7, 14 days). To do this, the techniques based on PFGE-Southern can be utilized.

It is expected that the signal(s) for AAVS1-specific integration in  $\Delta$ Ad.AAV/rep infected human cells increases during the first days after infection and then remains constant over time.

In a separate study, the integration status of vector DNA (analyzed by PFGE or PCR) and the number of integrated copies (analyzed by Southern Blot) with the expression level of  $\beta$ -galactosidase in single clones transduced with  $\beta$ -Gal hybrid vectors ( $\Delta$ Ad.AAV-BG,  $\Delta$ Ad.AAV-BG+Rep, or  $\Delta$ Ad.AAV-BG-rep) can be correlated. Together with data obtained in the studies described in the Specification, this allows assessment of whether transcriptional silencing is associated with site-specific vector integration into the AAVS1 site.

It is not clear a priori whether the specific Rep function for vector rescue, concatemerization, and integration can efficiently occur in non-S-phase or non-dividing cells. To test whether  $\Delta$ Ad.AAV<sup>fix</sup>,  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV-rep/+Rep vectors can integrate into non-dividing cells, transduction studies in cell cycle arrested cell cultures can be performed as described earlier.



## DISCUSSION

5 The establishment of stable cell lines expressing Rep 68/78 at detectable levels is not possible, which is probably due to rep mediated cytotoxicity. Therefore, it is not possible to perform long-term transduction studies (e.g. G418 selection or studies in single clones) in combination with ectopic rep expression. Moreover, due to the inhibitory effect of rep on adenovirus replication, it is currently not possible to generate adenoviral vectors expressing rep under the RSV or PGK promoter.

10

Taken together, this indicates that co-expressed Rep may stimulate site-specific transgene integration.

### 15 E. A Detailed Study Of Transduction/Integration Of Hybrid Vectors In Erythroid Cell Lines:

In order to improve transduction and integration frequencies of the hybrid vectors into erythroid cell lines, a detailed study comparing various hybrid vectors have to be carried out as described below. The transduction studies are performed in K562 cells which is  
20 considered to be an adequate model to study gene transfer vehicles into erythroid cells (Floch, V., et al., 1997, *Blood Cells, Mol. and Diseases*.23, 69-87). The optimal vectors should be able to integrate into the cellular genome with a high frequency, determined by Pulse field gel electrophoresis (PFGE) and Southern blot as described in Example 4. In addition, the results from the following studies will serve to evaluate whether a given  
25 hybrid vector needs to be modified for site-specific integration in the host genome.

#### Sequencing of integration junctions:

30 The ultimate proof for vector integration is the sequencing of junctions between SNori vector DNA and chromosomal DNA. Furthermore, this clarifies the question whether the AAV ITRs represent the substrate for integration. Specifically, DNA from clones with

known  $\Delta$ Ad.AAVSNori integration structure (analyzed by Southern Blot) digested with EcoRI, which does not cut within the SNori cassette. The resulting fragments are circularized and transformed into a specific *E. coli* strain (according to the protocol described by Rutledge and Russell (Rutledge, E. A. et al., 1997, *Journal of Virology*, 5 71:8429-36)). Kanamycin resistant bacterial clones should contain the integrated SNori cassette. Flanking chromosomal DNA in rescued plasmids can be sequenced with primers specific to the transgene.

To confirm vector integration in a small number of transduced cells, genomic DNA is 10 extracted and digested with EcoRI. EcoRI fragments are ligated to linkers containing a specific primer binding site and are then digested with NotI, religated and propagated in *E. coli*. Plasmid DNA from a representative number of bacterial clones is sequenced to determine the vector/chromosomal DNA junctions.

15 *Dose dependent toxicity:*

In order to test that the transduction frequency is dose-dependent and  $\Delta$ Ad.AAV vectors, which are devoid of all adenoviral genes, could be used to infect cells at higher doses with less cytotoxicity than first generation adenovirus, K562 cells are infected with 20 different MOIs ( $1-10^8$ ) of  $\Delta$ Ad.AAVBG and the first generation vector Ad.AAVBG (which contains the same  $\beta$ -Gal expression cassette). At day 4 post-infection, the total number of cells, the percentage of viable cells (based on trypan blue exclusion) and the percentage of X-Gal positive cells are counted. A fraction of infected cells are quantified for  $\beta$ -Gal expression using the Galacto-Light kit. The level of transgene expression is 25 expected to be comparable between the two vectors. K562 cells are predicted to tolerate higher doses of  $\Delta$ Ad.AAVBG better than Ad.AAVBG which express viral genes.

*Integration frequency with and without G418 selection:*

30 In order to investigate the integration frequency of the different vectors and to confirm that AAV ITRs present in double-stranded adenoviral DNA genomes can mediate vector

integration with a frequency comparable to rAAV vectors. integration studies are performed based on the formation of G418 resistant colonies with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.SNoriITR, and rAAVSNori after infection with  $2 \times 10^5$  and  $2 \times 10^6$  genomes per cell (Fig. 8). After infection, cells are plated in 96 well plates under limiting dilution and selected with G418 to estimate the frequency of formation of G418 resistant colonies. Another set of cells is plated without G418. A representative number of clones (w/ and w/o G418 selection) are expanded to  $>10^6$  cells (after 3-4 weeks of culture) and analyzed for the presence of viral DNA by Southern Blot as well as PFGE analysis to discriminate between episomal vector DNA and vector genomes stably associated with chromosomal DNA. This allows us to estimate the integration frequency of the different vectors, to assess the effect of G418 selection on integration, and to consider position effects on neo expression in calculating the total integration frequency. Integrated vector copies with a frequency of at least  $1 \times 10^{-4}$  is predicted only for  $\Delta$ Ad.AAVSNori and rAAVSNori. The total number of colonies may be lower in both the first generation vectors, Ad.AAVSNoriITR and Ad.SNori, due to the toxic effects of expressed adenoviral proteins; however, a higher integration frequency is predicted for the vector containing the AAV ITRs (Ad.AAVSNoriITR).

#### Kinetics of integration:

Compared to rAAV, the double-stranded nature of entering  $\Delta$ Ad.AAV genomes provides more protection against degradation. Furthermore, the synthesis of transcriptionally active double-stranded intermediates from single-stranded genomes, which is considered a limiting step in rAAV transduction, is not required in  $\Delta$ Ad.AAV transduction. Thus, the lag phase between infection and expression seen with rAAV vectors, which is causally linked to double-strand synthesis/integration may be shorter or absent in infections with  $\Delta$ Ad.AAV vectors. Furthermore, it was demonstrated earlier that a 9kb mini-adenoviral genome packaged into adenoviral particles is only short lived and completely degraded by day 3 post-infusion. In contrast, transduction with the 5.5kb  $\Delta$ Ad.AAV1 (Figure 1) genome allows for long-term expression, suggesting that either

AAV ITRs can stabilize the viral genome as an episome until it is integrated or integration occurs shortly after infection.

The status of vector DNA can be examined in K562 cells at different time points after infection with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.AAVSNoriITR, or rAAVSNori (MOI  $2 \times 10^5$ ). Infected cells are harvested at 1 hour, 5 hours, 1 day, 3, 7, and 14 days after infection and chromosomal DNA is analyzed by PFGE followed by hybridization with a transgene specific probe. This technique allows us to distinguish between episomal vector DNA, which appears as a distinct 5.0kb band and integrated DNA. Furthermore, extra chromosomal high-molecular weight vector concatemers can be detected. In the case of random integration, after digestion of chromosomal DNA with I-CeuI or PI-SceI, vector-specific signals in the range of 1-2mb should be seen. The intensity of episomal and integrated vector signal is quantified for each time point using phosphorimager analysis. This gives information about the kinetics of hybrid vector integration in a population of infected K562 cells and the intracellular stability of hybrid vector genomes.

Structure of integrated vector DNA and integration junctions with chromosomal DNA:

$\Delta$ Ad.AAV1 integrates as concatemer/s randomly into host DNA as shown previously. How many vector copies are present in one concatemer and whether the extent and the kinetics of tandem-formation are dose dependent still remain unclear. Another unanswered question is how  $\Delta$ Ad.AAV integrates: whether one or both ITRs are involved, whether the integrated ITRs are still intact, and whether adenoviral sequences integrate as well. These issues are important for the strategy to include rep genes into the hybrid vector genome. Moreover, if intact AAV ITRs are present within integrated vector copies, helper virus (adenovirus or HSV) infection in vivo may mobilize the integrated AAV-ITR vector cassette and affect stability of transgene expression.

To answer these questions, K562 cells can be infected with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.AAVSNoriITR, and rAAVSNori at MOIs  $2 \times 10^5$ ,  $2 \times 10^6$ , or  $2 \times 10^7$  genomes per cell. Infected cells are plated in 96 well plates in the presence or the absence of G418. The

latter is included because G418 may cause amplification of integrated vector DNA (Rutledge, E. A. et al., 1997, *Journal of Virology*, 71:8429-36). Genomic DNA from isolated clones can be analyzed by regular Southern Blot as described in the Examples Section to confirm the presence of vector concatemers and calculate the number of integrated vector copies. More informative is the sequencing of integrated vector copies and their junctions with chromosomal DNA. The structure of integration junctions can be delineated using the role of AAV ITRs in vector integration and the extent of insertional mutagenesis after transduction. This data provides information about the potential risks of hybrid vector used in clinical trials.

Transduction of cell cycle arrested cells:

The ultimate target for the hybrid vectors described in the Specification are quiescent hematopoietic stem cells. We hypothesize that the double-stranded nature of  $\Delta$ Ad.AAV genomes and specific nuclear import mechanisms may allow for the transduction of non-dividing cells. This is in part supported by the transduction studies with  $\Delta$ Ad.AAV1 in quiescent hepatocytes in vivo. To confirm this data, primary fibroblasts can be forced to enter the  $G_0$  phase by serum/growth factor starvation before infection with the hybrid vectors according to a protocol described by Russell (Russell, D. et al., 1995, *PNAS*, 92:5719-23). Cells are maintained for three days after infection under serum/growth factor deprivation. At this time point, genomic DNA is isolated and analyzed for integration events by PFGE in comparison with growing cells. Another series of integration studies can be carried out on K562 cells arrested in the  $G_1/S$  phase of the cell cycle with aphidicolin (added 1 day before and maintained several days after infection with hybrid vectors - depending on the integration kinetics studies described earlier). To investigate whether DNA damaging agents increase the transduction frequency of hybrid vectors, cell-cycle-arrested K562 cells or primary fibroblasts can be treated with cisplatin or  $^3H$ -thymidine prior to virus infection according to a protocol described by Alexander and Russell (Alexander, I. E. et al., 1994, *J. Virol.*, 68:8282-87; Russell, D. et al., 1995, *PNAS*, 92:5719-23). Furthermore, the effect of chromosomal DNA decondensation on the transduction efficiency of hybrid vectors can be studied in arrested

cells after treatment with puromycin, staurosporin, Hoechst 3328, distamycin, or vandate.

#### **F. Improvements in $\Delta$ Ad.AAV Production and Purification**

5 To inhibit packaging of full-length genomes a modified form of I-Sce I, a yeast mitochondrial intron-endonuclease with a non-palindromic 18-bp recognition sequence is expressed in 293 cells. Constitutive expression of this enzyme in mammalian cells is not toxic, possibly due to either the lack of I-SceI sites in the genome or sufficient repair of  
10 them (Rouet P. et al, 1994, *PNAS*, 91:6064-8). The yeast I-Sce I is modified with an SV40 T-antigen nuclear localization signal and an optimal Kozak sequence to enhance its functionality in mammalian cells (Rouet P. et al, 1994, *PNAS*, 91:6064-8). For another yeast endonuclease it was shown that a recognition site within an transduced Ad genome was efficiently (30% of all transduced genomes) when expressed in human A549 cells.  
15 Importantly, the expression of E4 ORF6 and ORF3 expressed from the transduced Ad genome inhibited double-strand break repair mediated by the endonuclease (Nicolas, A. L. et al, 2000, *Virology*, 266:211-24). This is consistent with the observations by others where these E4 proteins prevent concatemerization of the viral genome (Boyer, J. et al, 1999, *Virology*, 263:307-12). Based on this, packaging of full-length virus containing a  
20 I-SceI recognition site is reduced in 293 cells constitutively expressing I-Sce I. The 18mer I-Sce site is inserted into the E3 region of the Ad.IR vectors. These vectors are generated and amplified in 293 cells followed by a large-scale infection of 293 cells expressing I-SceI. Alternatively, an expression cassette for the endonuclease XhoI is inserted into the E3 region of Ad.IR or Ad.AAV vectors. The XhoI gene will be  
25 modified for optimal function in mammalian cells. Vectors expressing XhoI are generated and amplified in 293 cells expressing the Xho I isoschizomer PaeR 7 methyltransferase (PMT) (Nelson, J. E. et al, 1997, *J. Virol.*, 71:8902-7), which mediates the addition of a methyl group onto the N6 position of the adenine base of Xho I sites, CTCGAG. This protects the viral and cellular genome from XhoI cleavage. Methylated  
30 Ad vectors are produced at high titers.  $\Delta$ Ad.AAV vectors are then obtained by large-scale infection of 293 cells with the Ad.AAV-XhoI vectors. At this stage the viral

genome is not methylated and is digested at the XhoI sites. XhoI sites present within the transgene cassette are deleted by site-directed mutagenesis without altering the amino acids sequence. (XhoI is accumulated only at late stages in virus replication and should act only upon a large part of Ad DNA when replication is completed. In addition, ultracentrifugation optimizes the separation between  $\Delta$ Ad.IR and  $\Delta$ Ad.IR particles (Blague, C. et al., 2000, *Blood*, 95:820-8).

## EXAMPLE II

### 10 MODIFIED FIBER PROTEIN

#### A. Test the Infectivity of Different Human or Animal Serotype on Human Bone Marrow Cells.

15 Since the amino acid sequence of the fiber knob region varies considerably among the ~50 known serotypes, it is thought that different adenovirus serotypes bind to different cellular receptor proteins or use different entry mechanisms (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14; Defer, M., et al., 1993, *J. of*  
20 *Virology*, 64, 3661-3673). Although most adenoviruses contain RGD motifs in the penton base proteins, there are a number of serotypes (e.g. Ad 40, 41) without this conserved sequence. These types may use integrin  $\alpha$ v-independent pathways for virus internalization (Davison, A.J., et al., 1993, *J. Mol. Biol.*, 234, 1308-1316; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14). To test whether other Ad serotypes can  
25 infect stem cell subpopulation present in human bone marrow, studies with a series of different human Ad serotypes and animal viruses can be performed (see Table II). As a means to verify efficient transduction with Ad serotypes, viral DNA is tagged before infection and the presence of viral genomes in the nuclei of transduced cells is investigated. Furthermore, whether viral DNA is replicated in transduced cells can be  
30 analyzed as indirect proof for early viral gene expression. A direct detection of expressed viral proteins is impossible due to the unavailability of antibodies against all the serotypes included in this study. Simultaneously with the infection assay, transduced

human bone marrow cells can be analyzed for morphological and immunohistochemical features characteristic of HSC or progenitor subpopulations. For retargeting, serotypes which are able to infect CD34+ subsets of bone marrow cells at the lowest MOI are selected. As the next step, the fiber gene is PCR-cloned from serotypes with potential  
5 HSC/CD34+ tropism and inserted into standard shuttle plasmids for Ad5 vector generation replacing the Ad5 fiber gene using an E. coli recombination system (Figure 9).

## METHODS

10

### Cells and viruses:

HeLa (human cervix carcinoma, ATCC CCL-2.2), CHO (chinese hamster ovary, ATCC CCL-61), K562 (human hematopoietic, ATCC 45506), HEp-2 (human larynx carcinoma,  
15 ATCC CCL-23), 293 (human embryonic kidney, Microbix, Toronto Canada) cells were maintained in DMEM, 10% FCS, 2 mM glutamine, and Pen/Strep. Culture media for CHO cells was supplemented with 200µM asparagine and 200µM proline. Human CD34+-enriched bone marrow cells were purified from peripheral blood after mobilization using MiniMACS VS<sup>+</sup> separation columns (Miltenyi Biotec, Auburn, CA),  
20 according to the manufacturer's instructions. Aliquots were stored in liquid nitrogen. Sixteen hours before the experiment, cells were recovered from the frozen stock and incubated overnight in IMDM media, supplemented with 20% FCS, 10<sup>-4</sup> M β-mercaptoethanol, 100 µg/ml DNaseI, 2 mM glutamine, 10 U/ml IL-3, and 50 ng/ml stem cell factor (SCF) or 2 ng/ml thrombopoietin (Tpo). The purity of CD34+ preparations  
25 was verified by flow cytometry and was consistently greater than 90%.

### Flow cytometry:

Adherent cells (CHO, HeLa) grown in non-tissue culture treated 10 cm dishes (Falcon,  
30 Franklin Lakes, NJ) were detached by treatment with 1mM EDTA and washed three times with wash buffer (WB), consisting of PBS supplemented with 1% FCS. Cells



grown in suspension (K562, CD34+) were washed three times with WB. After washing, cells were resuspended in WB at  $2 \times 10^6$  cells/ml.  $2 \times 10^5$  cells were incubated in WB for 1 h at 37°C with monoclonal antibodies specific for  $\alpha_v$ -integrins [L230, ATCC: HB-8448, (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795) (1/30 final dilution),  
5 CAR [RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652) (1/400 final dilution)], or BrdU [(Amersham, Arlington Heights, IL) (1/100 final dilution)]. Subsequently, cells were washed with WB, and incubated with fluorescein isothiocyanate (FITC)-labeled horse anti-mouse IgG antibodies [(Vector Labs., Burlingame, CA) (1/100 final dilution)] or phycoerythrin (PE)-  
10 labeled goat anti-mouse IgG antibodies [(Calbiochem, La Jolla, CA) 1:100 dilution] for 30 min at 4°C. After incubation with secondary antibodies, cells were washed two times with WB and  $10^4$  cells per sample were analyzed in duplicate by flow cytometry.

For the analysis of CD34 and c-kit expression on transduced CD34+-cells and for  
15 fluorescent activated cell sorting (FACS), purified human CD34+ cells were incubated with phycoerythrin(PE)-conjugated anti-CD34 monoclonal antibodies (Becton-Dickinson Immunocytometry Systems, San Jose, CA) or with PE-labeled anti-CD117 (c-kit) monoclonal antibodies (MAb 95C3, Immunotech, Beckman Coulter, Marseille, France) according to the manufacturer's protocol followed by flow cytometry analysis. All  
20 analyses and sortings were performed on a FACStar Plus flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with 488nm argon and 633 nm HeNe lasers. For analysis of c-kit expression and FACS purification of CD34+/c-kit+ cells, SCF was not added to the media during culturing of CD34+ cells.

## 25 RESULTS

### CAR/ $\alpha_v$ -integrin expression on test cells:

It is generally accepted that CD34+ cells possess bone marrow repopulating activity.  
30 Therefore, we used human CD34+ cells as the target for our studies towards identifying Ad serotypes with HSC tropism and constructing new viral vectors. Studies were

performed on mobilized, CD34-positive, peripheral blood cells from one donor under conditions which are known to retain CD34+ cells in a quiescent stage (Leitner, A., et al. 1996. *Br.J.Haematol.* 92:255-262; Roberts, A. W., Metcalf, D. 1995. *Blood.* 86:1600-1605). More than 90% of purified cells were CD34 positive by flow cytometry.

5 Furthermore, we included into our Ad tropism studies the cell line K562, which is considered to be an adequate model for studying gene transfer into human hematopoietic cells (McGuckin, et al. 1996. *British Journal of Haematology.* 95:457-460). HeLa cells, which are readily infectible by Ad5, and CHO cells, which are refractory to Ad5 infection (Antoniou, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9), were used as positive and

10 negative control cell lines, respectively.

For Ad5, both, binding to the primary receptor and to  $\alpha_3\beta_5$  and  $\alpha_v\beta_5$  integrins are important for high efficiency infection of target cells. The expression of CAR and  $\alpha_v$  integrins on test cells was analyzed by flow cytometry using monoclonal antibodies

15 against CAR (RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652)) and  $\alpha_v$  integrins (L230 (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621)) (Figure 10). As expected, nearly all HeLa cells expressed high levels of CAR and  $\alpha_v$ -integrins, whereas CHO cells lacked significant

20 CAR and  $\alpha_v$ -integrin expression. Fifteen and 77% of K562 cells expressed CAR and  $\alpha_v$ -integrins, respectively. Only ~6% of the CD34+ cells used in our studies expressed CAR and 17% were positive for  $\alpha_v$ -integrins. Notably, the preparation of CD34+ cells represents a mixture of different cell types. The absent or low expression of primary and secondary Ad5 receptors on non-cycling human CD34+ cells is in agreement with

25 previous reports (Huang, S., et al. 1996. *J. Virology.* 70:4502-4508; Neering, S. J., et al. 1996.. *Blood.* 88:1147-1155; Tomko, R. P., et al.. 1997. *Proc. Natl. Acad. Sci. USA.* 94:3352-3356).

Infection assay using wild-type Ad5 and K562 cells:

30 The presence of viral DNA in the nucleus of infected cells is an indirect means to demonstrate efficient virus binding, internalization, and nuclear import. Nuclear

localization of the viral genome is a prerequisite for transgene transcription and integration. Two techniques are utilized to tag viral DNA for in situ analysis. To optimize the infection assay, wild-type Ad5 virus and K562 cells which are permissive for Ad5 infection can be used. The first protocol (Challberg, S.S. and Ketner, S. 1981, *Virology* 114, 196-209), is based on  $^{32}\text{P}$ -labeling of viral DNA. During amplification of wild-type Ad5 and A549 cells,  $^{32}\text{P}$ -phosphate ( $40\mu\text{Ci/ml}$ ) is added to phosphate-free medium. After development of CPE,  $^{32}\text{P}$ -tagged virus is harvested, banded in CsCl gradients, and titered on HeLa cells according to standard protocols. To simulate the conditions for infection of human bone marrow cells, K562 cells are incubated in suspension with a MOI of 1, 10, or 100 of  $^{32}\text{P}$ -Ad5 for 2, 4, 6, or 8 hours under agitation at  $37^\circ\text{C}$ . This covers the time period necessary for adsorption, internalization, and nuclear import. After washing, cells are fixed either transferred to microscopy slides using cytopsin or embedded in paraffin and sectioned (according to protocols from VECTOR labs, Burlingame, CA). The latter has the potential advantage that multiple consecutive sections ( $5\mu\text{m}$ ) of the same cell can be analyzed by different methods (e.g. for  $^{32}\text{P}$  tagged viral DNA, for specific histological staining, for immunofluorescence), which allows for correlating infection with a particular cell type present in the bone marrow. Cells are incubated in a Kodak NTB-2 photo emulsion for autoradiography. The exposure time can be optimized to minimize background or non-nuclear localized signals. A dose and time dependent appearance of nuclear silver grains is expected under the optimized conditions. Since  $^{32}\text{P}$ -phosphate can label viral proteins as well, a cytoplasmic background signal might appear. To facilitate detection, immunofluorescence with HSC specific antibodies on sections can be performed. As an alternative method, a BrdU-labeling technique for viral DNA can be used (Lieber, A., et al. 1999. *J Virol* 73:9314-24; Lieber, A. et al., 1996, *Journal of Virology*, 70:8944-60). In this case, different amounts of BrdU are added to the A549 culture medium during wtAd5 virus propagation. BrdU labeled viral DNA can be detected with monoclonal antibodies specific to BrdU. The signal can be enhanced using layers of species-specific polyclonal antibodies in combination with biotin/avidin and a fluorescent marker. BrdU tagged viral DNA can be detected on cytopsin of bone marrow cells together with cell surface markers by double or triple immunofluorescence.

## DISCUSSION

- The interaction of selected Ad serotypes with CD34+ cells was tested. As a result of this screening we constructed a first-generation, Ad5-based vector whose fiber was substituted with the fiber derived from Ad35. We demonstrated that this capsid modification allowed for efficient viral transduction of potential HSCs by the corresponding chimeric Ad vectors.
- All tropism and transduction studies were performed with non-cycling CD34+ cells, which are thought to include HSCs. The quiescent stage of CD34+ cells purified from mobilized blood is important because induction of cell proliferation is associated with a loss of the ability to reconstitute hematopoiesis and with changes in the spectrum of cellular receptors (Becker, P. S., et al. 1999. *Exp. Hematol.* 27:533-541). It is known that treatment of hematopoietic cells with cytokines or growth factors changes the expression of specific integrins including  $\alpha_v$ -integrins, which would ultimately alter the susceptibility of cells to Ad infection or may effect viability of infected cells (Gonzalez, R., et al. 1999. *Gene Therapy*. 6:314-320; Huang, S., et al. 1995. *J. Virology*. 69:2257-2263). Another fact that complicates the interpretation of transduction studies is the extraordinary heterogeneity of CD34+ cells in regards to morphology and function.

### B. Screening different adenoviruses to establish tropism to HSC.

- The ATCC provides more than 70 different human or animal adenoviruses (see Appendix I). A collection of 15 human serotypes and 6 animal adenoviruses (see Table II) are selected based on the following criteria: (i) availability of the complete genome sequence or fiber sequence from the NIH gene bank (ii) CAR receptor usage absent or unknown, (iii) different subgroups, and (iv) moderate or low tumorigenicity (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia).
- However, any serotype shown in the Appendix hereto can be used for the invention described. Animal viruses are included in the infectivity assay because this may provide a

means to circumvent the pre-existing humoral immunity against human Ad5 fiber, which represents a critical obstacle for clinical trials with Ad vectors.

## METHODS

5

### Viruses:

The following human adenovirus serotypes were purchased from the ATCC: 3 (VR-3), 4 (VR1081), 5 (VR-5), 9 (VR1086), 35 (VR-716) and 41(VR-930). Adenovirus No. VR-  
10 716 was purchased from ATCC labeled as serotype 34, however it was found to be serotype 35 upon sequencing of the fiber region. For amplification, the corresponding Ads were infected onto HeLa, 293, or HEp-2 cells under conditions that prevented cross-contamination. Virus was banded in CsCl gradients, dialyzed and stored in aliquots as described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).  
15 Plaque titering was performed as follows: Confluent 293 cells plated in 6-well plates were incubated for 24 hours with virus in a total volume of 1ml. Two weeks after infection, plaques were counted on cultures overlayed with 1% agarose/MEM/10% FCS.0

### 20 EM studies:

CsCl-banded Ad stocks were thawed and diluted with 0.5% glutaraldehyde. Grids were prepared as described earlier (Mittereder, N., et al. 1996. *J. Virology*. 70:7498-7509). After staining with 2% methylamine tungstate (Nanoprobes, Stony Brook, NY), the  
25 carbon-coated grids were evaluated and photomicrographed with a Phillips 410 electron microscope, operated at 80 kV (final magnification 85,000x). For each particular Ad serotype, the number of morphologically deficient viral particles per 100 was counted in five random fields.

### 30 RESULTS

Electron microscopy:

Little is known about the stability of particles from serotypes other than Ad5. Since the intactness of viral particles was crucial for comparative interaction studies, virions from the serotypes specified above were analyzed by electron microscopy (EM). EM studies of negative contrast stained Ad suspensions demonstrated that the percentage of defective particles (loss of icosahedral shape or luminal staining) did not exceed 5% indicating that serotype preparations had comparable qualities. Representative EM photographs are shown for Ads 5, 9, and 35 (Figure 11).

Serotype screening:

It is thought that different Ad serotypes bind to different cellular receptor proteins and use different entry mechanisms (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Mathias, P., et al. 1994. *Journal of Virology*. 68:6811-6814). A set of human adenoviruses was obtained from the ATCC to be tested for tropism to CD34+ cells. These included serotypes 3, 4, 5, 9, 35, and 41 representing different subtypes (Table 1). We believed that these serotypes would use different cellular attachment and internalization strategies due to differing lengths of fiber shafts (Chroboczek, J., et al. 1995. *Adenovirus fiber*, p. 163-200. In a. P. B. W. Doerfler (ed.), *The molecular repertoire of adenoviruses*, vol. 1. Springer Verlag, Berlin; Roelvink, P. W., et al. 1998. *J. Virology*. 72:7909-7915), the presence or absence of RGD motifs within the penton base, and differing tissue tropism. The relatively little characterized Ad35 was selected because it was found in immunocompromised hosts, particularly in bone marrow recipients (Flomenberg, P., et al. 1994. *Journal of Infectious Diseases*. 169:775-781; Flomenberg, P. R., et al. 1987. *Journal of Infectious Diseases*. 155:1127-1134; Shields, A. F., et al. 1985 *New England Journal of Medicine*. 312:529-533). The latter observations prompted us to believe that bone marrow cells are among the natural reservoirs for Ad35.

**TABLE II****Human and animal adenoviruses with potential interest for the invention**

| Adeno virus | Human/ Group B            | Human/ Group D   | Human/ Group F | Avian     | Bovine | Canine | Ovine | Swine | Mouse |
|-------------|---------------------------|------------------|----------------|-----------|--------|--------|-------|-------|-------|
| Sero type   | <u>3,7,11,16,21,34,35</u> | 8,15,17,19,28,37 | <u>40, 41</u>  | CELO, EDS | 3      | 1,2    | 5     | 4     | 1     |

The underlined serotypes use CAR independent pathways for cell entry.

- 5 For amplification, the corresponding adenovirus stocks can be infected onto HeLa or A549 cells such that at a given time only one virus type is handled in a separate laminar flow hood and cultured in Hepa-filtered bottles, preferentially in separate CO<sub>2</sub> incubators to avoid cross-contamination. During propagation, viral DNA is tagged using one of the techniques described earlier. Viral DNA can be isolated from purified particles. The
- 10 XhoI restriction pattern is analyzed for methylated and unmethylated viral DNA by Southern blot using the full genome of the corresponding virus type as a radioactive probe.

### DISCUSSION

15

- Although it was reported earlier by slot-blot assay that fiber knobs derived from 2, 9, 4, and 41L can bind to CAR (Roelvink, P. W., et al. 1998. *J. Virology*. 72:7909-7915), it is not clear whether this binding occurs with an affinity that is physiologically relevant and whether this would confer cell entry. Furthermore, as shown for the Ad5 interaction
- 20 between the penton and integrins, a secondary receptor is required to induce virus internalization. We demonstrated that different serotypes interacted differently with the K562 or CD34+ target cells. Ad5, Ad4, and Ad41 were not able to efficiently attach to and be internalized by K562 and CD34+ cells. Although Ad4 belongs to a separate subgroup (E), it is thought that Ad4 represents a natural hybrid between subgroup B and
- 25 C viruses with a fiber related to Ad5 (Gruber, W. C., et al. 1993. *Virology*. 196:603-611).

Therefore, it was not surprising that Ad4 has binding properties similar to Ad5. The subgroup F serotype Ad41 has been shown to contain distinct fibers, a long shafted and a short-shafted fiber allowing for different cell entry pathways (Tiemessen, C. T., Kidd, A.H. 1995. *J. Gen. Virol.* 76:481-497). The Ad41 penton base does not contain RGD motifs suggesting that this virus may use  $\alpha_v$ -integrin independent pathways for cell entry. However, these features did not improve interaction with CD34+ cells. Ad9, Ad3, and Ad35 did interact with CD34+ cells more efficiently than Ad5. Out of all the serotypes tested, Ad35 demonstrated the most efficient attachment and internalization with K562 and CD34+ cells. Although the short-shafted Ad9 can bind to CAR, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Therefore, the low level of  $\alpha_v$ -integrin expression on certain subset/s of CD34+ cells may account for the observed susceptibility to Ad9.

#### C. Attachment and Internalization of the Ad serotypes to K562 and CD34+ cells.

#### METHODS

##### Labeling of Ads with [<sup>3</sup>H]-methyl thymidine:

Serotypes were labeled with [<sup>3</sup>H]-methyl thymidine as described in detail elsewhere (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Briefly,  $5 \times 10^7$  HeLa or 293 cells were grown in 175 sq. cm flasks with 15 ml DMEM/10% FCS and infected with wild type adenovirus at a MOI of 50 or higher. Twelve hours post-infection, 1 mCi of [<sup>3</sup>H]-methyl thymidine (Amersham, Arlington Heights, IL) was added to the media and cells were further incubated at 37°C until complete CPE was observed. Then, cells were harvested, pelleted, washed once with cold PBS, and resuspended in 5 ml PBS. Virus was released from the cells by four freeze-thaw cycles. Cell debris was removed by centrifugation and viral material was subjected to ultracentrifugation in CsCl gradients and subsequent dialysis as previously described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology.* 70:8944-8960). Virus purification and dialysis removed unincorporated radioactivity. Wild type Ad particle concentrations were determined



spectrophotometrically by measuring the OD<sub>260</sub>, utilizing the extinction coefficient for wild-type Ad5  $\epsilon_{260} = 9.09 \times 10^{13}$  OD ml cm virion<sup>-1</sup> (Maizel, J. V., et al. 1968. *Virology*. 36:115-125). The virion specific radioactivity was measured by a liquid scintillation counter and was always in the range of  $1 \times 10^5$  to  $1 \times 10^4$  cpm per virion. For selected  
5 variants, the fiber gene was PCR amplified and sequenced to ensure identity and the absence of cross-contamination.

Viral DNA tagged with methylase and test for replication by genomic Southern blots:

10 To ultimately confirm transduction, a protocol to detect adenoviral replication in infected cells can be established. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. A site-specific methylation strategy is utilized to monitor viral DNA replication within infected cells (Nelson, J. et al., 1997, *Journal of Virology*, 71:8902-07). Methylation marked adenovirus can be produced by the addition of a  
15 methyl group onto the N6 position of the adenine base of XhoI sites, CTCGAG, by propagation of the virus in HeLa or A549 cells expressing the XhoI isoschizomer PaeR7 methyltransferase (PMT) (Kwoh, T.J., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 7713-7717). It is known that methylation does not affect vector production but does prevent cleavage by XhoI. Loss of methylation through viral replication restores XhoI cleavage  
20 and can be detected by Southern blots of genomic DNA from infected cells in comparison to native, non-methylated, viral genomes.

Attachment and internalization assays:

25 These studies were performed based on a protocol published elsewhere (Wickham, T. J., et al. 1993. *Cell*. 73:309-319). In preliminary experiments, we found that labeled Ad5 virions reached equilibrium in attachment to HeLa cells after 45 min at 4°C with an MOI of 400 pfu per cell. For attachment studies,  $3.5 \times 10^5$  cells were incubated for one hour on ice with equal amounts of [<sup>3</sup>H]-labeled adenovirus OD particles equivalent to an MOI of  
30 400 pfu/cell for Ad5 in 100 µl of ice-cold adhesion buffer (Dulbecco's modified Eagle's medium supplemented with 2 mM MgCl<sub>2</sub>, 1% BSA, and 20 mM HEPES). Next, the cells

were pelleted by centrifugation for 4 min at 1000 x g and washed two times with 0.5 ml ice-cold PBS. After the last wash, the cells were pelleted at 1500 x g, the supernatant was removed, and the cell-associated radioactivity was determined by a scintillation counter. The number of viral particles bound per cell was calculated using the virion specific radioactivity and the number of cells. To determine the fraction of internalized [<sup>3</sup>H]-labeled adenoviral particles, cells were incubated on ice for one hour with the corresponding virus, washed with PBS as described above, resuspended in 100 µl adhesion buffer, and then incubated at 37°C for 30 min. Following this incubation, cells were diluted 3-fold with cold 0.05% trypsin-0.5mM EDTA solution and incubated at 37°C for an additional 5-10 min. This treatment removed 99% of attached radioactivity. Finally, the cells were pelleted at 1500 x g for 5 min, the supernatant was removed, and the protease-resistant counts per minute were measured. This protocol minimizes the possibility that the internalization data were affected by receptor recycling (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795). Nonspecific binding of Ad particles to cells on ice was determined in the presence of 100-fold excess of unlabeled virus. This value routinely represented less than 0.1 % of viral load.

## RESULTS

### Attachment of Ad particles to target cells and internalization:

The selected serotypes were metabolically labeled with [<sup>3</sup>H]-thymidine, which is incorporated into viral DNA during replication. Adsorption and internalization can be experimentally dissociated by taking advantage of the observation that at low temperature (0-4°C) only virus cell attachment occurs, whereas internalization requires incubation at higher temperatures. The number of particles adsorbed or internalized per cell was calculated using the virion-specific radioactivity and used to quantify interaction of Ads 3, 4, 5, 9, 35, and 41 with CD34+, K562, HeLa and CHO cells (Figure 12). The serotypes varied significantly in their ability to attach to and to be internalized by the different cell lines. For Ad5, the degree of attachment to the cell lines tested correlated with the level of CAR expression. In CHO cells, which were previously shown to be refractory to Ad5

infection, the level of attachment and internalization was about 50-70 viral particles per cell. This number was hereafter assumed negative in terms of susceptibility of a given cell type for Ad5. Interaction of the other serotypes with CHO cells was not significantly higher indicating that corresponding receptor/s were absent on CHO cells. All serotypes tested interacted with HeLa cells; with Ad3 and Ad35 being the most efficient variants. The presence of distinct Ad3 and Ad5 receptors on HeLa cells was demonstrated previously (Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857). Ads 4, 5, and 41 did not bind to K562 cells. In contrast, Ad9 as well as the members of subgroup B, Ad3 and Ad35, efficiently interacted with K562 cells with Ad35 having the highest number of adsorbed and internalized particles. Compared to Ad5, about 25 times more Ad35 particles were attached and three-fourth of these were internalized by K562 cells. Viral interactions with CD34+ cells were generally weaker. Among the serotypes tested, only Ad9 and Ad35 were significantly internalized by non-cycling CD34+ cells. Internalization of Ad9 and Ad35 was, respectively, four and eight times more efficient than for Ad5 particles. The number of Ad35 virions internalized by CD34+ cells was almost half of that seen for Ad5 in HeLa cells, which can be readily infected with Ad5 based vectors.

Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into HeLa, 293, and CHO cells:

HeLa and 293 cells expressing high level of primary and secondary receptors for human adenoviruses are used as a positive control for virus attachment and internalization. As a negative control CHO cells are used. CHO cells do not express the primary adenoviral receptor at a detectable level, and are therefore refractory for adenoviral infection. For attachment studies, these adherent cell lines are detached from 10 cm dishes with PBS-EDTA solution (without Ca<sup>2+</sup> and Mg<sup>2+</sup>), washed three times with ice-cold PBS, resuspended in adhesion buffer, and incubated with viruses as described above in the Examples section. As expected, all adenoviral serotypes tested are efficiently attached to and internalized into HeLa cells (Table III) (Figure 13). Adenoviruses serotypes 3, 5, 35, 41, but not 9, are efficiently attached to and internalized by 293 cells. In contrast, poor

attachment and internalization of most adenovirus serotypes are observed with CHO cells. The level of attachment on CHO is about 50-70 virus particles per cell for adenoviruses serotypes 5 and 41, 115 virus particles per cell for adenovirus type 3 and about 180 particles per cell for adenovirus serotypes 9 and 35. For further analysis, numbers >300 viral particles per cell are assumed as positive and <70 viral particles per cell as negative in terms of susceptibility of a particular cell line for efficient adenoviral transduction.

TABLE III

10 Comparative analysis of attachment and internalization of Ad5 and Ad9 to cell lines, expressing different amounts of CAR and  $\alpha\beta$ integrins.

| Cell line    | CAR expression | $\alpha\beta$ -integrin expression | Ad9 (attached/internalized) | Ad5 (attached/internalized) |
|--------------|----------------|------------------------------------|-----------------------------|-----------------------------|
| HeLa         | ++             | ++                                 | 426/370                     | 550/500                     |
| CHO          | -              | ++                                 | 300/300                     | 70/50                       |
| 293          | ++             | ++                                 | 20/20                       | 1950/1750                   |
| Y79          | +++            | -                                  | 190/140                     | 1200/1100                   |
| K562         | -              | +                                  | 320/230                     | 60/50                       |
| Erythrocytes | ?              | ?                                  | 420/-                       | 68/-                        |

15 Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into human CD34+ bone marrow cells and K562 erythroleukemia cell line:

Previous studies showed that the human erythroleukemia cell line K562 can be transduced with Ad5-based adenoviral vectors at very high MOIs. As shown in Figure 14, only about 60 viral particles per cell of adenovirus serotype 5 are attached to and even fewer particles are internalized into these cells at a MOI of 400. In contrast to Ad5, about 320 viral particles per cell of Ad9 and about 1500 viral particles per cell of Ad35 are

attached to and about two-thirds of them are internalized into K562 cells (Figure 14B). Human unstimulated CD34+-enriched bone marrow cells obtained from frozen stocks are incubated overnight in growth medium without cytokine stimulation. The next day, the number of viable cells is calculated. For attachment studies, cells are washed three times  
5 with ice-cold PBS, resuspended in adhesion buffer and incubated with adenoviruses. Among the adenoviral serotypes tested, only adenovirus particle of Ad9 (about 150 viral particles per cell) and Ad35 (about 320 viral particles per cell) are able to attach to unstimulated CD34+ cells on the level, compared to Ad 5 (only 60 viral particles per cell). Four-fifths of these virus particles are able to be internalized by the cells.  
10 Interestingly, upon stimulation of CD34+ cells with GM-CSF and EPO/TPO for two weeks, attachment and internalization of Ad9 viral particles are significantly increased (up to 300 particles per cell). At the same time, the transient stimulation of cells with GM-CSF for two days could not increase the level of viral attachment to the cells.

15 Based on the above finding that Ad35 serotype is able to attach and internalize into CD34+ cells most efficiently among several serotypes tested, serotype Ad35 was selected for further studies. As described in Appendix II, a chimeric vector (Ad5 GFP/F35) containing the short-shafted Ad35 fiber sequence in an Ad5 capsid was able to target a broad spectrum of CD34+ cells in a CAR/integrin independent manner.

## 20 *DISCUSSION*

In summary, from all the serotypes tested, Ad9, Ad3, and Ad35 demonstrated the most efficient attachment to and internalization with K562 and CD34+ cells. Based on  
25 adsorption/internalization data, Ad9 and Ad35 as representatives for subgroups D and B were selected for further tropism studies.

### **D. Characterization of Ad vector replication in K562 and CD34+ cells.**

30 Comparative analysis of Ad5, and Ad9 and Ad34 to infect and to replicate in 293, K562 and CD34+ cells. The ability of the Ad9 fiber knob domain to recognize the same

primary receptor on the cell surface as Ad5 with comparable affinity was described earlier. Thus, the finding that Ad9 viral particles can only poorly attach to 293 cells is rather unexpected. In order to find out how the attachment and internalization data reflect the biological activity of adenoviruses of different serotypes, the stocks of Ad5, Ad9 and Ad35 are characterized in more detail by electron microscopy, plaque assay on 293 cells, and quantitative replication assay in K562 and CD34+ cells.

## METHODS

### Quantitative replication assay:

1x10<sup>5</sup> CD34+ or K562 cells were infected in 100µl of growth media with different MOIs of Ad5, 9, or 35 which had been amplified in 293 cells, expressing the XhoI DNA methyltransferase isoshizomer PaeR7 (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). After 2 hours of incubation at 37°C, the cells were centrifuged at 1000 x g for 5 min, the virus-containing medium was removed, the cells were resuspended in 100µl of fresh media, and then they were incubated at 37°C until harvesting. At 16 hours post-infection for K562 cells, or 36 h post-infection for CD34+ cells, 5 µg of pBS (Stratagene, La Jolla, CA) plasmid DNA was added as a carrier which could also be used as a loading control. Genomic DNA was extracted as described previously (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). One-fourth of purified cellular DNA (equivalent to 2.5 x 10<sup>4</sup> cells) was digested with HindIII, XhoI, or with HindIII and XhoI together at 37°C overnight and subsequently separated in a 1% agarose gel followed by Southern blot with chimeric Ad5/9 or Ad5/35 DNA probes. The chimeric probes, containing sequences of Ad5 and Ad9 (Ad 5/9) or Ad5 and Ad35 (Ad 5/35), were generated by a two-step PCR amplification using Pfu-Turbo DNA polymerase (Stratagene, La Jolla, CA) and viral DNA from purified particles as a template. The following primers were used for PCR (Ad5 sequences and nucleotide numbers are underlined): Ad5F1 – (nt: 32775-32805) 5'-GCC CAA GAA TAA AGA ATC GTT TGT GTT ATG-3'; Ad5R1 – (nt: 33651-33621) 5'-AGC TGG TCT AGA ATG GTG GTG GAT GGC GCC A-3'; chimeric Ad5/9F – (nt: 31150-31177, nt: 181-208 ) 5'-AAT

GGG TTT CAA GAG AGT CCC CCT GGA GTC CTG TCA CTC AAA CTA GCT  
GAC CCA -3'; chimeric Ad5/9R - (nt: 32805-32775, nt:1149-1113) 5'-CAT AAC  
ACA AAC GAT TCT TTA TTC TTG GGC TTC ATT CTT GGG CGA TAT AGG  
AAA AGG-3; chimeric Ad5/35F - (nt: 31150-31177, nt: 132-159) 5'-AAT GGG TTT  
5 CAA GAG AGT CCC CCT GGA GTT CTT ACT TTA AAA TGT TTA ACC CCA-3',  
chimeric Ad5/35R (nt: 32805-32775, nt: 991-958) 5'-CAT AAC ACA AAC GAT TCT  
TTA TTC TTG GGC ATT TTA GTT GTC GTC TTC TGT AAT GTA AG-3'.  
Nucleotide numbers are given according to the sequences obtained from the NCBI  
GenBank (accession No. M73260 / M29978 for Ad5, X74659 for Ad9, and U10272 for  
10 Ad35). After the first amplification, the 968 bp-long Ad9, a 859 bp-long Ad35 DNA  
fragments corresponding to the fiber genes, and a 876 bp-long Ad5 fragment  
corresponding to the Ad5 E4 region (located immediately downstream of Ad5 fiber gene)  
were purified by agarose gel electrophoresis. To generate chimeric DNA probes,  
amplified Ad5 DNA was mixed with Ad9 or Ad35 fragments obtained during the first  
15 step of PCR, and subjected to a second PCR amplification using Ad5/9F or Ad5/35F  
primers and the Ad5R1 primer. The resulting Ad5/9 or Ad5/35 chimeric DNA fragments  
(see Fig. 15) were purified and their concentrations were measured  
spectrophotometrically. Corresponding chimeric DNA fragments were loaded as  
concentration standards on agarose gels or labeled with [<sup>32</sup>P]-dCTP and used as probes  
20 for Southern analysis. The number of viral genomes per DNA sample was calculated  
after quantitative Phospho-imager analysis. In preliminary experiments, no preferential  
hybridization of chimeric DNA probes to DNA of any particular viral serotype was  
detected.

## 25 RESULTS

### Replication of selected serotypes in K562 and CD34+ cells:

Adsorption/internalization studies do not ultimately prove viral transduction, a process  
30 often defined as gene transfer that allows for viral or heterologous gene expression in  
host cells. Intracellular trafficking, including endosomal lysis, transport to the nucleus,

and nuclear import of the viral genome, depends on structural capsid proteins and thus, varies between different serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065). We believed that analysis of viral gene expression would be a means to verify successful nuclear import of viral genomes and that this would be a good criterion for selection of serotype/s able to efficiently infect our target cells. To do this, we used a protocol, which allows for the detection of Ad replication in infected cells. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. We utilized a site-specific methylation strategy to monitor viral DNA replication within infected cells (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). Methylated Ad serotypes were produced by the addition of a methyl group onto the N6 position of the adenine base of Xho I sites, CTCGAG, during propagation of the viruses in 293 cells expressing the Xho I isoschizomer PaeR 7 methyltransferase (PMT) (Kwoh, T. J., et al. 1986. *Proc.Natl.Acad.Sci. USA*. 83:7713-7717) (293 PMTcells). Loss of methylation through viral replication restores Xho I cleavage and can be detected by Southern blots of Xho I-digested genomic DNA from infected cells.

Ad replication studies were performed in K562 and CD34+ cells with Ad9 and Ad35, in comparison to Ad5. For replication studies, the infectious titer (in pfu/ml) and genome titer (in genomes per ml) were determined (by plaque assay on 293 cells or by quantitative Southern blot, respectively) for methylated and unmethylated Ad5, Ad9, and Ad35 (Table 2). The ratio of pfu to genome titer was comparable for methylated and unmethylated virus demonstrating that DNA methylation had not altered transduction properties. About 85% of (Ad5, 9, and 35) virus used for infection was methylated as calculated based on the intensity of fragments specific for methylated and non-methylated viral DNA present in the viral load (Fig. 15). The numbers of genomes detected after adsorption (1 hour, 0°C) or internalization (2 hours 37°C) correlated well with studies shown in Fig. 12. Ad9 and Ad35 interacted more efficiently than Ad5 with K562 and CD34+ cells. Dose-dependent replication studies in K562 and CD34+ cells were performed with the same genome numbers of Ad5, 9, and 35 (Fig. 15). The replication rate was measured based on the ratio of methylated to demethylated viral



DNA after infection with different MOIs (2100, 420, and 105 genomes per cell). In K562 cells, efficient replication (100% conversion from methylated to unmethylated DNA) was detected for Ad5 at MOI  $\geq 2100$ , for Ad9 at MOI  $\geq 420$ , and for Ad35 at MOI  $\geq 105$ . This demonstrated that Ad35 transduced K562 cells with the highest efficiency.

5 In CD34+ cells, the replication rate was 100% for Ad5 and 31% for Ad9 after infection with MOI 420. Although methylated Ad35 viral DNA was present in CD34+ cells, viral replication was undetectable for Ad35. In summary, while viral replication studies in K562 cells confirmed data obtained for Ad5, 9, and 35 adsorption and internalization, there was a discrepancy between earlier results and the poor replication of Ad9 and,

10 particularly, Ad35 in CD34+ cells. As outlined later, replication analysis in heterogeneous cell populations, like CD34+ cells, may not allow for definitive conclusions on tropism of a particular serotype.

Taking all the screening data together, Ad9 and Ad35 emerged as the variants with the

15 strongest tropism for K562 and CD34+ cells. It is thought that Ad9 can bind to CAR, however, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology*. 70:7614-7621). This entry strategy may be not optimal for efficient infection of CD34+ cells as only less than 17% of them express  $\alpha_v$ -integrins (Fig. 10). Therefore, we decided to concentrate on Ad35 as a source for heterologous fiber to be used for

20 construction of a chimeric vector based on an Ad5 backbone.

**TABLE IV**

Results from the infectivity assay which determines the optical particle-to-PFU (OPU/PFU) ratio using 293 cells

25

| Virus | OPU (A260)            | PFU                   | OPU/PFU ratio |
|-------|-----------------------|-----------------------|---------------|
| Ad5   | $1.4 \times 10^{12}$  | $1.06 \times 10^{11}$ | 13            |
| Ad9   | $4.61 \times 10^{11}$ | $2.6 \times 10^8$     | 1773          |

**DISCUSSION**

Viral replication studies in K562 cells confirmed the data obtained for Ad5, 9, and 35 adsorption and internalization. However, there was a discrepancy between the interaction data and the replication data in CD34+ cells where Ad9 replicated only poorly and no replication was seen for Ad35. Ad replication is only initiated upon the production of a critical threshold of early viral proteins, which in turn, is directly dependent on the number of viral genomes present in the nuclei of infected cells. Therefore, the outcome of replication studies may be affected by the rate of nuclear import of viral genomes, by the activity of viral promoters, and/or the intracellular stability of viral DNA/RNA. These parameters may vary, on one hand, between different subsets of CD34+, and/or, on the other hand, between different Ad serotypes. In conclusion, the viral replication analyses performed with different Ad serotypes in CD34+ cells may not predict the actual transduction properties of chimeric vectors based on Ad5 backbone. This implies that attempts to produce gene transfer vectors based on Ad genomes other than Ad5 should be exercised with caution.

Recently, an Ad serotype screening strategy was used to identify variants with tropism for primary fetal rat CNS cortex cells or human umbilical vein endothelial cells. The optimal serotype (Ad17) was selected based on immunohistochemistry for hexon production 48 hours after infection (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540). However, this approach is problematic because, at least in our hands, antibodies developed against Ad5 hexon did not cross-react with other serotypes. Also, hexon is expressed only after onset of replication. As outlined above, the kinetics of intracellular trafficking, viral gene expression, and replication significantly vary between serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065).

In addition to being the most efficient serotype in terms of interaction with CD34+ cells, Ad35 is also interesting because it interacts with receptor/s different from the Ad5 and Ad3. Ad35 and Ad5GFP/F35 attachment was not inhibited by Ad5 or anti-CAR antibodies suggesting that Ad35 binding was CAR independent. First, Ad5 did not compete with Ad35 and Ad5GFP/F35 during internalization and infection indicating that

$\alpha_5\beta_3/5$  integrins are not involved in viral entry. Second, function-blocking antibodies against  $\alpha_v$ -integrins did not compete with Ad35 and Ad5GFP/F35 for internalization into K562 cells, whereas these antibodies did inhibit Ad5 internalization. And third, in contrast to Ad5 based vectors, GFP expression after infection with Ad5GFP/F35 was not restricted to  $\alpha_v$ -integrin-expressing CD34+ cells. From these facts, we conclude that infection with Ad35 and the chimeric Ad5GFP/F35 vector does not involve  $\alpha_v$ -integrins. In this context, the presence or absence of RGD motifs within Ad35 penton base remains to be determined by sequencing the corresponding genome region. Cross-competition assays demonstrated that Ad35 and Ad5GFP/F35 bind to a receptor that is different from the Ad3 receptor. Although Ad3 and 35 belong to the same subgroup, they have been divided into two DNA homology clusters, B1 and B2; the amino acids composing their fibers are only 60% homologous. Furthermore, the target tissues for both viruses are different; Ad3 can cause acute respiratory infections, whereas Ad35 is associated with kidney infection (Horwitz, M. S. 1996. *Adenoviruses*, p. 2149-2171. In B. N. Fields, Knipe, D.M., Howley, P.M. (ed.), *Virology*, vol. 2. Lippincott-Raven Publishers Inc., Philadelphia). Therefore, it was not surprising to see that Ad3 and Ad35 recognize different receptors.

In conclusion, Ad35 and the chimeric vector enter the cells by a CAR- and  $\alpha_v$ -integrin independent pathway. We believe that Ad35 and the chimeric vector binds primarily to its fiber receptor and that this interaction is sufficient to trigger internalization. On the other hand, Ad35 internalization may involve cellular proteins other than  $\alpha_v$ -integrins. These membrane proteins can overlap with those for Ad3 internalization and represent  $\beta_2$  integrins, which protrude more from the cell surface than  $\alpha_v$ -integrins (Huang, S., et al. 1996. *J. Virology*. 70:4502-4508).

According to EM studies of negative contrast-stained adenoviral suspensions, the percentage of deficient particles for all adenoviral serotypes tested does not exceed 5%. However, plaque assays reveal that the ability to form plaques in 293 cells is significantly different for tested serotypes. The optical particle-to-PFU (OPU/PFU) ratio obtained is 13 for Ad5, which is in good agreement with the previously estimated ratio for this

adenoviral serotype. Importantly, this ratio is about three times higher for adenovirus serotype 35 and more than 150-fold higher for adenovirus serotype 9. Furthermore, quantitative Southern blot using chimeric Ad5/9 and Ad5/35 DNA probes is used to determine the ratio between the genome and transducing titer. This study confirms the data obtained by plaque assay. Quantitative replication assay of these adenoviruses in K562 and CD34+ cells also confirms the ability of Ad9 and Ad34 to more efficiently attach to these cell types. The replication of viral genomes is observed for Ad9 and Ad34 at lower MOIs of infection, compared to Ad5. In conclusion, the data obtained for different serotypes in attachment and internalization are in good agreement with the infectivity data in target cells.

**E. Attachment and internalization of different adenoviral serotypes into primary dendritic cells, JAWSII, MCF-7 and REVC cells.**

As a proof of principle, the serotype screening strategy can be employed for other important target cells which are refractory to Ad5 infection.

***RESULTS***

RECV cells are endothelial cells which have to be targeted for approaches that are aimed to gene therapy of restenosis, atherosclerosis, inflammation etc. MCF-7 cells are breast cancer cells isolated from liver metastases which are important targets for tumor gene therapy. The human adenovirus serotypes 3, 5, 9, 35 and 41 are tested to see whether they can attach to and can be internalized by mouse primary dendritic cells, JAWSII cells, MCF-7-human breast cancer cells and REVC endothelial cells. None of the adenoviral serotype tested can efficiently attach to primary dendritic cells. Adenovirus serotype 3 is able to efficiently attach to REVC endothelial cells (about 400 virus particles per cell are attached and about 300 are internalized). In comparison, only 50 Ad5 particles are able to attach to and even fewer are internalized in these REVC. The human breast cancer cells (MCF-7) are previously shown to be refractory to Ad5

infection at low MOIs. However, Ad3 and more efficiently, Ad35 attach to and internalize into MCF-7 cells.

## DISCUSSION

5

The data presented herein indicate that different human adenovirus serotypes recognize different cellular receptors and can therefore infect cell types that are refractory to Ad5 infection. There are adenoviral serotypes that can more efficiently attach and internalize than Ad5 for human CD34+ cells, REVC, K562 and MCF-7 cells. This finding provides  
10 a basis for the construction of chimeric adenoviral vectors which are Ad5 vectors containing receptor ligands derived from other serotypes.

### F. Infection studies on primary human bone marrow cells.

15

Since established erythroleukemic cell lines do not represent an adequate model for the ultimate hematopoietic stem cell that has to be targeted in patients in order to achieve long-term reconstitution with genetically modified cells, normal primary human bone marrow cells are used for the initial infection/retargeting studies.

20

## RESULTS

In a first set of tropism studies with different Ad serotypes, whole bone marrow cell suspensions can be used without preselection. This is advantageous because the tropism of various adenovirus serotypes or genetically retargeted vectors can be analyzed on a  
25 broad spectrum of progenitor subpopulations representing myeloid, erythroid, megakaryocytic, lymphoid, dendritic, and monocytic lineages. For short term (< 5 hours) infection studies, bone marrow suspensions can be cultured in IMDM supplemented with 10% FCS,  $\beta$ -mercaptoethanol, and 10u/ml IL-3 for ensuring cell viability.

30

Mononucleated cell assays:

Mononucleated bone marrow cells can be incubated with MOI 1, 10, 100, or 1000 pfu/cell of the various adenovirus types for a short time. Paraffin sections or cytopins of  
5 infected bone marrow cells can be analyzed for nuclear-localized, labeled viral DNA. BrdU labeling can be visualized by immunofluorescence with anti-BrdU antibodies; <sup>32</sup>P-tagged viral DNA can be detected by incubation with photo-emulsion. In addition, the same cell material can be analyzed for morphology after specific histo-staining (e.g. Wright-, Hemo3 staining). If required, commercially available antibodies can be used to  
10 specific cell surface markers conjugated directly to different fluorochromes (FITC (green), TRIT., RPE, (red), RPE-Cy5, AMCA (blue)) to completely characterize infected bone marrow subpopulations. Colocalization of BrdU-labeled viral DNA (e.g. as FITC signal) with membrane markers signifying infection of specific cell types can be demonstrated; for example, potential stem cells/early progenitors (CD34<sup>+</sup>, CD38<sup>+</sup>),  
15 megakaryocytes (CD41a<sup>+</sup>), erythroid cells (glycophorin A<sup>+</sup>), dendritic cells (CD1a<sup>+</sup>), monocytes (CD14<sup>+</sup>), or myeloid cells (CD15<sup>+</sup>), etc. The morphological analysis of infected bone marrow subsets gives a first information whether specific adenovirus serotypes can target primitive cell types.

20 *DISCUSSION*

Since the different wild-type adenoviruses do not express a uniform marker gene and do not integrate and since detection of tagged viral DNA cannot be done on live cells, it is not possible, at this point, to characterize infected cells for clonogenic or repopulation  
25 capacities. Therefore, adenovirus serotypes for retargeting studies are selected, based on their ability to infect in vitro purified CD34<sup>+</sup> cells at low MOIs. This subset of bone marrow cells is known to contain long-term reconstituting cells. Infection studies with different adenovirus serotypes can be repeated on purified CD34<sup>+</sup> cells (cultured in IMDM +10% FCS,  $\beta$ -mercaptoethanol, and 10 units/ml IL-3) as described above.  
30 Purification of CD34<sup>+</sup> cells can be performed by direct immunoadherence on anti-CD34 monoclonal antibody-coated plates or on MiniMacs columns as described by

Papayannopoulou (Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69; Papayannopoulou, T. et al., 1993, *Blood*, 81:229). The purity of isolated CD34+ cells ranges routinely from 80-95%. Analogous infection studies can be repeated with selected adenovirus types on CD34+/CD38- subsets.

5

To confirm productive infection purified CD34+ cells can be infected with selected (methylase-tagged) serotypes and analyze viral DNA replication. Cultures of purified human bone marrow CD34+ cells can be used for the transduction and integration studies as a model for HSCs.

10

It was recently demonstrated that HSC activity does exist in CD34-negative human bone marrow subsets (Bathia, M. et al., 1998, *Nature Medicine*, 4:1038-45; Osawa, M., et al., 1996, *Science*, 273:242-5; Goodell, M. et al., 1997, *Nature Medicine*, 3:1337-45; Zanjani, E. D. et al., 1998, *Exp. Hematology*, 26:353-60). Lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup> cells can be tested in the retargeting and transduction studies in combination with repopulation assays in SCID-NOD mice.

15

#### G. Cloning and insertion of the fiber gene.

#### 20 *METHODS*

##### PCR-cloning of the corresponding fiber gene and insertion into Ad5 based shuttle plasmids instead of the endogenous AD5 fiber:

25 One or several adenoviruses with tropism to CD34+ or other HSC containing population is selected for further studies described herein. The complete coding region for fiber varies between 1-2kb, depending on the virus type. The fiber encoding sequences can be obtained by PCR with Pfu polymerase from viral DNA isolated from purified particles of the selected virus types. The corresponding primers can be designed based on the fiber sequences available from the EMBL gene bank. The PCR products are cloned as PacI-BalI fragment into pCD4 (Figure 10), a shuttle vector for recombination of RecA+ E.

30

coli. In pCD4, the heterologous fiber gene is flanked on both sides with Ad5 sequences, which are homologous to regions directly adjacent to the fiber reading frame in Ad5. As an Ad5 (shuttle vector) derived template for recombination, pCD1, a pBHG 10 (Microbix, Toronto, Canada) derivative can be used. The recombination procedure is performed according to a protocol routinely used for recombinant adenovirus generation (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Routinely, 90% of the resulting plasmids are accurately recombined. The junctions between the heterologous fiber (X) and Ad5 sequences can be sequenced to confirm the accuracy of recombination. The resulting plasmid is named pAd5fiberX (pAd5<sup>fx</sup>). The resulting product is used to generate pAd5<sup>fx</sup>-based Ad.AAV containing the heterologous fiber gene.

Construction of chimeric Ad vectors:

For transduction studies, two Ad vectors were constructed: Ad5GFP and Ad5GFP/F35, containing a chimeric Ad5/35 fiber gene. Both adenoviral vectors contained a 2.3kb, CMV promoter driven EGFP gene [derived from pEGFP-1, (Clontech, Palo Alto, CA)] inserted into the E3 region of Ad5. The EGFP expression cassette was cloned between Ad5 sequences 25,191-28,191 and 30,818-32,507 into a shuttle plasmid, which contained the E3 deletion described for pBHG10 (Microbix, Toronto, Canada). The resulting plasmid was named pAdGFP. For the chimeric vector, the Ad5 fiber gene in pAdGFP was substituted by an Ad5/35 chimeric fiber gene generated by the two-step PCR protocol outlined above. In the first PCR step, three DNA fragments corresponding to i) the Ad5 fiber 5'-nontranslated region and the first 132 bp of the fiber tail domain (nt 30,818-31,174), ii) the Ad35 shaft and knob domains (nt 132-991), and iii) the Ad5 E4 region including the Ad5 fiber polyadenylation signal (nt 32,775-33,651) were amplified by Pfu-Turbo DNA polymerase. The following primers were used: for the Ad5 tail, Ad5F-2 (nt 30,798-30,825) 5'-CGC GAT ATC GAT TGG ATC CAT TAA CTA-3' and Ad5R-2 (nt 31,174-31,153) 5'-CAG GGG GAC TCT CTT GAA ACC CAT T-3'; for the Ad35 shaft and knob, primers Ad5/35F and Ad5/35R (see above); for the Ad5E4 and polyA, primers Ad5F-1 and Ad5R-1 (see above). After 10 PCR cycles, the products were purified by agarose gel electrophoresis, combined, and then subjected to a second PCR



with primers Ad5F-2 and Ad5R-1. The resulting 2115 bp-long chimeric fiber gene contained the Ad5 tail and the Ad35 shaft and knob domains. This product was used as a substitute for the SalI/XbaI Ad5 fiber gene containing fragment in pAdGFP. The resulting plasmid was named pAdGFP/F35. To generate full-length E1/E3 vector  
5 genomes, pAdGFP and pAdGFP/F35 were inserted in pAdHM4 (Mizuguchi, H., Kay, M.A. 1998. *Human Gene Therapy*. 9:2577-2583) by recombination in *E.coli* (Chartier, C., E. et al. 1996. *Journal of Virology*. 70:4805-4810). To do this, the RecA+ *E.coli* strain BJ5183 was co-transformed with pAdHM4 linearized by SrfI mixed with the XbaI fragments containing the GFP genes, the Ad5 or Ad5/35 fiber genes, and the Ad5  
10 homology regions. The resulting recombinants were analyzed by restriction analysis. Correct recombinants were amplified in *E.coli* HB101 and purified by double CsCl gradient banding. The plasmids were named pAd5GFP and pAd5GFP/F35. The correct structure of the Ad5/35 chimeric fiber gene was confirmed by endonuclease digestion and sequencing part of pAd5GFP/F35. To produce the corresponding viruses, pAd5GFP and  
15 pAd5GFP/F35 were digested with PacI to release the viral genomes and transfected onto 293 cells as described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). Plaques developed 7 to 10 days post-transfection in overlaid cultures. Recombinant viruses were propagated in 293 cells and purified by standard methods described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

20

Hemagglutination assay:

Twenty-five microliters of serial dilutions of Ad5, Ad35, or chimeric Ad5GFP/F35 virions in McIlvaine-NaCl buffer (0.1 M citric acid, 0.2 M Na<sub>2</sub>HPO<sub>4</sub> [pH 7.2], diluted  
25 1:50 with 0.87% NaCl) were loaded onto 96 well plates. To each dilution, 25 µl of a 1% suspension of monkey erythrocytes (in McIlvaine-NaCl buffer) was added. The sedimentation pattern was determined after incubation for 1 hour at 37°C. All tests were performed in quadruplicates in at least two independent experiments.

30

Southern blot:

Extraction of genomic DNA, labeling of DNA fragments and hybridization were performed as described earlier (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

RESULTSConstruction/Characterization of chimeric fiber:

Previously, it has been shown that exchanging the fiber knob was sufficient to alter the tropism of chimeric Ad vectors (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V., et al. 1998. *J. Virology*. 72:1844-1852; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). As outlined above, the length of the fiber shaft may critically determine the entry strategy of a particular serotype. Therefore, we decided to replace not only the Ad5 fiber knob but also the shaft. The chimeric Ad5/35 fiber contained the Ad5 tail (amino acid: 1-44) necessary for interaction with the Ad5 penton base linked to 279 amino acids from Ad35 including the shaft with 7  $\beta$ -sheets and the knob (Fig. 16A). The endogenous Ad5 fiber polyA signal was used to terminate transcription of the chimeric fiber RNA. The combination of the Ad5 capsid including the RGD motif containing penton base with a short-shafted fiber could be risky because the natural distance between the fiber knob and the RGD motifs was disturbed. The Ad5 fiber was substituted by the chimeric fiber sequences based on an E1/E3 deleted Ad vector. This vector carried a CMV promoter-GFP reporter gene cassette inserted into the E3 region. The corresponding chimeric virus (Ad5GFP/F35) was produced in 293 cells at a titer of  $>2 \times 10^{12}$  genomes per ml. For comparison, an E1/E3 deleted Ad vector containing the original Ad5 fiber gene and the GFP expression cassette was generated (Ad5GFP). The titer and the ratio of physical to infectious particles was similar between Ad5GFP and Ad5GFP/F35 indicating that the fiber modification did not significantly alter the stability and/or growth properties of the chimeric vector. The correctness of the fiber modification was confirmed by restriction analysis of the Ad5GFP/F35 viral genome followed by

Southern blot hybridization (Fig. 16B), direct sequencing of the fiber-coding region, and a functional test for hemagglutination (HA) of monkey erythrocytes. The agglutination of erythrocytes is fiber knob-mediated; it is known that Ad5 does not agglutinate monkey erythrocytes whereas Ad35 efficiently does (Pring-Akerblom, P., et al. 1998. *J. Virology*. 72:2297-2304). In HA tests, Ad5GFP/F35 agglutinated monkey erythrocytes with the same efficiency as Ad35 at dilutions of up to 1:512. In contrast, no hemagglutination was observed with equivalent Ad5 dilutions. This clearly confirmed the functional activity of the chimeric Ad5/35 fiber incorporated into Ad5 capsid.

10 Generation of chimeric adenoviral vectors (Ad.AAV<sup>h</sup>) with heterologous fiber molecules: Adenoviruses with chimeric Ad5-Ad3 fiber are viable and can be produced at high titers (Krasnykh, V., et al., 1996, *J. of Virology*, 70, 6839-6846; Stevenson, S. C. et al., 1997, *J. Virology*, 71:4782-90). In order to test whether the fiber substitution described herein affects production or stability of adenoviruses, two E1-deleted first-  
15 generation, adenoviral vectors are produced with the AAV- $\beta$ gal cassette in 293 cells using standard protocols. The vector is generated by recombination of pAd.AAV-BG (Fig. 17) with pCD1 (containing the endogenous Ad5 fiber); the other vector (with heterologous fiber) is the recombination product of pAd.AAV  $\beta$ Gal and pAd5fiberX (pAd5<sup>h</sup>). Virus from single plaques is amplified on 293 cells. The production yield per  
20 293 cell can be determined by plaque-titering of 293 cell Lysates. It is anticipated that the fiber modification will not critically affect the stability of chimeric vectors. Finally, bone marrow cells can be infected with the retargeted vectors. Two days after infection, live-cell cytometry is performed for  $\beta$ -gal expression using as substrate Fluorescein di- $\beta$ -D-Galactopyranoside (FDG) (Cantwell, M.J. et al., 1996 *Blood* 88, 4676-4683; Neering, S. et al., 1996, *Blood*, 88:1147-55; Fiering, S. N. et al., 1991, *Cytometry*, 12:291; Mohler, W. et al., 1996, *PNAS*, 93:57) and the infected cells are characterized for morphology and surface markers. Before and during infection, bone marrow cells can be cultured in IMDM/FCS supplemented with thrombopoietin (Tpo), which supports the survival of HSC (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996,  
25 *Experimental Hematology*, 24:660-69). Alternatively, retargeted vectors can be  
30

generated with the AAV-GFP (green fluorescence protein) cassette and perform FACS analysis on transduced cells based on GFP and surface marker expression.

#### H. Competition studies of chimeric fiber protein Ad5/35.

5

##### Competition studies:

Cross-competition studies between Ad5, 35, and Ad5GFP/F35 (Fig. 18) for binding and internalization were performed in order to investigate in more detail the pathways which are used by the chimeric vector to infect target cells. Wild-type Ad35 and the chimeric vector Ad5GFP/F35 could recognize the same primary receptor as they competed with each other for the attachment to K562 cells (Fig. 19A, upper panel). This primary receptor is different from that used by Ad5, since neither Ad5 viral particles nor anti-CAR monoclonal antibodies (Fig. 19B, upper panel) were able to abrogate Ad35 or Ad5GFP/F35 binding. In competition studies for internalization, Ad35 and Ad5GFP/F35 competed with each other with equal efficiency. Ad5 and anti- $\alpha_v$ -integrin monoclonal antibodies (L230) (Figs. 19C, D; lower panel) did not inhibit internalization of Ad35 or the chimeric virus. To consolidate this data, K562 cells were infected with Ad5GFP and Ad5GFP/F35 after prior incubation of cells with anti-CAR or anti- $\alpha_v$ -integrins monoclonal antibodies followed by analysis of GFP-expressing cells. The transduction data mirror the results obtained in adsorption/internalization studies. In summary, this demonstrated that Ad35 and Ad5GFP/F35 use a CAR and  $\alpha_v$ -integrin-independent pathway for infection of K562 cells; the structural elements which account for these specific properties are located within the Ad35 fiber and can be transplanted into Ad5 by fiber substitution.

Ad3 can efficiently interact with K562 cells (Fig. 12), although Ad3 and Ad35 belong to the same subgroup (B), the homology between amino acid sequences of their fibers is only about 60%. Therefore, we decided to test whether Ad3 could compete with Ad35 and Ad5GFP/F35 for attachment and internalization (Fig. 20). These studies demonstrated that Ad35 binding was not inhibited by Ad3 indicating the use of different

receptors. Interestingly, Ad3 slightly inhibited attachment of Ad5GFP/F35 (Fig. 20A, left panel). In addition to binding to the receptor common for the Ad35 and Ad5GFP/F35 fiber, the chimeric capsid (e.g. the Ad5 penton RGD motifs) may also interact with a second cellular receptor that overlaps with elements involved in Ad3 binding. In cross-competition for internalization, pre-incubation of cells at 37°C with Ad35 and with chimeric virus significantly decreased internalization of [<sup>3</sup>H]-labeled Ad3 (Fig. 20D, right panel). In the reverse experiment, Ad3 as competitor decreased the level of internalization by 30% for both, Ad35 and the chimeric virus (Fig. 20B, right panel). As expected, Ad5 and Ad3 did not compete for adsorption or internalization. As shown before (Fig. 19B), anti-CAR and anti- $\alpha_v$ -integrin antibodies did not block Ad3 interaction with K562 cells. In summary, we concluded that Ad35 and Ad5GFP/F35 bind to receptor/s different from that of Ad3, although they can use common structural elements for internalization, which are different from  $\alpha_v$ -integrins.

15 *Infection studies with chimeric virus:*

It is established that Ad5GFP/F35 infected K562 cells by a CAR and  $\alpha_v$ -independent pathway. It is possible that this property allows for efficient transduction of non-cycling CD34+ cells, which express scarcely CAR and  $\alpha_v$ -integrins. To test this, the transduction properties of Ad5GFP and Ad5GFP/F35 vectors were analyzed on CD34+ cells, K562, and HeLa cells. Fig. 21 shows the percentage of transduced, GFP expressing cells depending on the MOI used for infection. Nearly 100% of HeLa cells were transduced with Ad5GFP and Ad5GFP/F35 at MOIs of  $\geq 25$ . More than 95% of the K562 cells were transduced with Ad5GFP/F35 at MOIs of  $\geq 100$ , whereas the transduction rate was significantly lower with Ad5 where it increased with the MOI reaching a plateau at ~70% GFP-positive cells after infection with an MOI of 400. Transduction of CD34+ cells was about three fold more efficient with Ad5GFP/F35 than with Ad5GFP at all MOIs analyzed. Interestingly, at higher MOIs, the transduction rate did not rise proportionally with the viral dose and soon reached a plateau indicating that in both cases only specific subset/s of CD34+ cells were permissive to infection. In order to characterize in more detail these specific, permissive subset/s, additional transduction

studies were performed. First, the percentage of GFP expressing cells was determined in CD34+ fractions that were stained for  $\alpha_v$ -integrins or CARs (Fig. 22). The low number of CAR positive CD34+ cells complicated accurate co-labeling studies, and there was no correlation between CAR expression and the proportion of transduced cells among CD34+ cells infected with Ad5GFP or Ad5GFP/F35. Interestingly, for Ad5GFP, 65% of all GFP expressing cells were positive for  $\alpha_v$ -integrins, whereas less than 22% of GFP positive cells infected with the chimeric virus stained positive for  $\beta_1$ -integrin expression. While only 17% of the whole CD34+ population expressed GFP after Ad5GFP infection, the percentage of GFP-expressing cells in the CD34+/ $\beta_1$ -integrins positive fraction was 50%. This indicates that Ad5GFP vector-mediated GFP expression was preferentially localized to  $\alpha_v$ -integrin positive CD34+ subsets, whereas after infection with the Ad5GFP/F35 vector, GFP was expressed in a broader spectrum of CD34+ cells with most of them being  $\alpha_v$ -integrin-negative.

Next, transduced cells were simultaneously analyzed for GFP as well as for CD34 and CD117 markers. As mentioned before, only about 90% of all cells used in our analysis were positive for CD34 at the time of infection, hence the multiparameter analysis for CD34 and GFP. A population of CD34+ cells is extraordinarily heterogeneous in morphology and stem cell capacity. The subpopulation of CD34+ and CD117+ cells resembles very primitive hematopoietic cells (*Ikuta, K., Weissman, I.L. 1992 Proc. Natl. Acad. Sci. USA. 89:1502-1506; Simmons, P. J., et al. 1994. Expl. Hematology. 22:157-165*). Fig. 23 summarizes the analyses of GFP expression in correlation with these specific stem cell markers. While 54% of cells infected with chimeric vector were positive for GFP and CD34+, only 25% of cells infected with Ad5GFP expressed the transgene and CD34+ marker (Fig. 23A, lower panel). More importantly, based on GFP expression, the chimeric virus transduced 80% of c-kit positive cells, whereas the Ad5-based vector transduced only 36% (Fig. 23A, middle panel). In an additional experiment, CD34+ cells were sorted for CD117 expression prior to infection with Ad5GFP or Ad5GFP/F35 and, 24 hours post-infection, GFP expression was analyzed in this specific fraction (Fig. 23B). This analysis revealed that the chimeric vectors transduced 4 fold more CD34+/CD117+ than the Ad5GFP vector.

In conclusion, these results demonstrated that the chimeric Ad5GFP/F35 vector was clearly superior to the Ad5GFP vector in targeting and transduction of CD34+ cells. Furthermore, the data suggest that the spectrum of CD34+ cell subsets permissive for Ad infection was significantly different for the chimeric vector than for the Ad5 vector.

Analysis of viral genomes within CD34+ cells infected with the Ad5 and chimeric vectors:

So far, the transduction rate of CD34+ cells was measured based on GFP expression after infection with Ad5GFP and Ad5GFP/F35. Considering the extraordinary heterogeneity of CD34+ cells in morphological and functional parameters, GFP may not be expressed in all cell types that were efficiently infected. Reasons for this include that the CMV promoter may not be active in all cell types or that the regulation of transgene expression could differ between subsets on a post-transcriptional or post-translational level. To test this, we quantified the number of intracellular (transduced) viral genomes within GFP positive and GFP negative fractions of CD34+ cells infected with Ad5GFP and Ad5GFP/F35. To do this, twenty-four hours after infection, CD34+ cells were sorted for GFP positive and GFP negative fractions, which were subsequently used to isolate genomic DNA together with transduced viral DNA. The number of viral genomes was determined by quantitative Southern blot as described for Fig. 15. Per GFP-positive CD34+ cell, about 270 copies of the Ad5GFP/F35 viral genome were detected. Interestingly, a remarkable 200 copies of the Ad5GFP/F35 viral genome were found per GFP-negative CD34+ cell (Fig. 24A and 25). This demonstrated that not all infected cells expressed GFP and implies that the actual transduction rate was higher than 54% (GFP-positive cells). We concluded that the CMV promoter was not active in all transduced CD34+ subsets. No Ad5GFP vector specific signal was detected within infected CD34+ (GFP positive or negative) fractions by Southern blot which had a detection limit of 14 viral genomes per cell. From this, we can conclude that the vector DNA concentration per transduced cell was at least 20 times higher for Ad5GFP/F35 than for Ad5GFP.

Ad5GFP DNA was only detectable in DNA samples from infected CD34+ cells by Southern blot after prior PCR amplification with vector specific primers (Fig. 24B and 25). This indicates that the replication deficient Ad5 vector is present but at a very low copy number, which may be limited by intracellular genome stability. Using the PCR-  
5 Southern detection method, Ad5 vector DNA was also detected in GFP negative cells, supporting that the CMV promoter may not have been the optimal choice for transduction studies. It is notable that studies by others on viral genome analyses after infection of CD34+ cells with Ad5 vectors were performed only after prior PCR amplification (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Neering, S. J., et al. 1996..  
10 *Blood*. 88:1147-1155).

### DISCUSSION

The chimeric Ad5GFP/F35 vector has binding and internalization properties similar to  
15 Ad35. Therefore, the fiber substitution was sufficient to swap cell tropism from Ad5 to Ad35. The Ad5GFP/F35 capsid chimera contained the short-shafted Ad35 fiber incorporated into an Ad5 capsid, instead of the naturally occurring long-shafted Ad5 fiber. During Ad5 infection, interaction between the penton base and integrins is required to induce viral internalization. For this interaction, the length of fiber shaft and  
20 the precise spatial arrangement of knob and RGD motifs are critical for the virus entry strategy. The natural spatial arrangement is disturbed when short-shafted heterologous fibers are inserted into the Ad5 capsid. Interestingly, the Ad5/35 capsid chimera allows for efficient infection, suggesting that the protruding RGD motives in the Ad5 penton base do not affect the interaction with the primary Ad35 receptor. So far, most of the  
25 chimeric viruses were generated by substituting only the Ad5 knob while maintaining the long Ad5 fiber shaft (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V. N., et al. 1996. *J. Virology*. 70:6839-6846; Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). The exception was an Ad5/7 chimeric virus (Gall, J., et al. 1996. *J. Virology*. 70:2116-2123), where the  
30 whole Ad5 fiber was substituted by the short-shafted Ad7 fiber. However, similar to the parental Ad5, the Ad5/7 chimera still required  $\alpha_v$ -integrins for infection.



This Ad5GFP/F35 chimera is the first demonstration that despite the presence of RGD motifs within the Ad5 penton, the chimeric virus uses cell entry pathways determined primarily by the receptor specificity of the short-shafted heterologous fiber. This does not  
5 exclude that interaction with a secondary receptor may increase binding affinity. The latter is supported by the observation that Ad35 and Ad5GFP/F35 slightly differed in their ability to compete with Ad5 or Ad3 for binding. It is possible that Ad5/35 attachment involves, in addition to the high affinity fiber binding, interaction between Ad5 capsid proteins (e.g. RGD motifs) and secondary receptor/s that overlap with those  
10 used by Ad3 and Ad5.

This data indicate that infection with Ad5-based vectors is restricted to a specific subset of CD34+ cells. The percentage of GFP expressing cells after Ad5GFP infection of CD34+ cells reached a plateau at MOIs higher than 100 indicating that only a limited  
15 fraction of CD34+ cells was permissive to Ad5. Also, strong replication of wild type Ad5 in infected CD34+ cells may be the result of preferential transduction of a specific subpopulation of CD34+ resulting in a expression of early viral genes at a level sufficient to initiate viral replication. The presence of a specific subpopulation of CD34+ cells permissive to Ad5-vector infection was suggested by others (Byk, T., et al. 1998. *Human*  
20 *Gene Therapy*. 9:2493-2502; Neering, S. J., et al. 1996.. *Blood*. 88:1147-1155). In the present report, we further characterized this subpopulation and demonstrated that Ad5-based vectors preferentially infected  $\alpha_v$ -integrin positive CD34+ cells. Integrins (including  $\alpha_v$ ) are thought to be important for homing and trafficking of transplanted hematopoietic cells, however little is known about the correlation between  $\alpha_v$ -integrin  
25 expression and the differentiation status of hematopoietic cells (Papayannopoulou, T., Craddock, C. 1997. *Acta Haematol.* 97:97-104; Roy, V., Verfaillie, C.M. 1999. *Exp. Hematol.* 27:302-312). There was no clear correlation between CAR and GFP expression suggesting that Ad5GFP may be able to use another membrane protein as a primary receptor. Alternatively, Ad5GFP transduction observed at an MOI of 200-400 could be  
30 the result of direct interaction between virus and  $\alpha_v$ -integrins triggering internalization, which may be the preferred pathway in the absence of CAR (Legrand, V., et al. 1999. *J.*

*Virology*. 73:907-919). Importantly, infection with the chimeric Ad5GFP/F35 vector was not restricted to the  $\alpha_v$ -positive CD34+ subpopulation.

Among CD34+ cells, the subpopulation of CD34+ and CD117+ cells resembles very primitive hematopoietic cells (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*. 89:1502-1506; Simmons, P. J., et al. 1994. *Expl. Hematology*. 22:157-165). The receptor for stem cell factor, CD117 (c-kit) belongs to a tyrosine kinase family. It was previously shown that c-kit+, CD34+ cord blood cells contain a high fraction (16%) of hematopoietic progenitors (Neu, S., et al. 1996. *Leukemia Research*. 20:960-971). Early in ontogeny 34+/CD117+ cells have long-term repopulating activity (Sanchez, M. J., et al. 1996. *Immunity*. 5:513-525). An average of 50-60% of CD34+ cells are reported to be CD117 positive (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*. 89:1502-1506; Neu, S., et al. 1996. *Leukemia Research*. 20:960-971; Simmons, P. J., et al. 1994. *Expl. Hematology*. 22:157-165). In our studies, the chimeric vector expressed GFP in 54% CD34+ cells and 80% of CD34+/c-kit+ cells. The actual viral transduction rate could be even higher because transduced Ad5GFP/F35 vector DNA was also found in GFP-negative fractions of infected cells. This indicates that the CMV promoter used to drive GFP expression in our vectors was not active in all transduced cells. We selected the CMV promoter for transgene expression based on published data demonstrating that PGK and CMV promoters allowed for efficient transgene expression in CD34 cells whereas the HTLV-I and RSV promoter were almost inactive (Byk, T., et al. 1998. *Human Gene Therapy*. 9:2493-2502; Case, S. S., et al. 1999. *Proc. Natl. Acad. Sci. USA*. 96:2988-2993). On the other hand, studies by Watanabe et al. (Watanabe, T., et al. 1996. *Blood*. 87:5032-5039) suggest that the CMV promoter is not active or rapidly silenced in certain CD34+ subsets. Our data underscore this observation. Considering retroviral transduction studies, the retroviral MLV promoter may have been a better candidate for transduction studies in hematopoietic cells (Bregni, M., et al. 1998. *Gene Therapy*. 5:465-472).

After having demonstrated that the Ad5GFP/F35 vector efficiently transduced cells carrying stem cell specific markers, the next logical step would be to perform colony

assays with pre-sorted GFP positive/negative cells. However, this assay is complicated by the fact that infection with first generation Ad vectors is cytotoxic and affects the formation and growth of progenitor colonies in MC-cultures (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Watanabe, T., et al. 1996. *Blood*. 87:5032-5039). This side effect is caused by the expression of Ad proteins within transduced cells (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960; Schiedner, G., et al. 1998. *Nature Genetics*. 18:180-183; Yang, Y., et al. 1994. *Proc. Natl. Acad. Sci. USA*. 91:4407-4411). Some of these proteins (e.g E4-orf4, pTP, or E3-11.6k) have pro-apoptotic activity (Langer, S. J., Schaak, J. 1996. *Virology*. 221:172-179; Lieber, A., et al. 1998. *J. Virology*. 72:9267-9277; Shtrichman, R., Kleinberger, T. 1998. *J. Virology*. 72:2975-2983; Tollefson, A. E., A et al. 1996 *J. Virology*. 70:2296-2306). Clearly, this would affect the outcome of transduction studies with Ad5GFP/F35, which allows for the efficient transfer of viral genomes into CD34+ cells implying significant expression of viral proteins. Moreover, recently published data indicate that short-term colony assay mostly measure mature progenitors and do not represent a rigorous test for transduction of potential stem cells.

A definitive demonstration that Ad5GFP/F35 based vectors can transduce HSC requires colony assays or preferably, repopulation assays in SCID-NOD mice. We can perform these studies with gutless vectors (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) and integrating  $\square$ Ad.AAV vectors devoid of all viral genes (Lieber, A., et al. 1999. *J Virol* 73:9314-24) generated based on Ad5GFP/F35 chimeric capsids. Alternatively, gutless, retargeted vectors could be used to transiently express a retroviral receptor on CD34+ cells to increase their susceptibility to infection with retroviral vectors based on an approach that we have published earlier (Lieber, A., et al.. 1995. *Human Gene Therapy*. 6:5-11).

Our finding that Ad5GFP/F35 can efficiently transduce hematopoietic cells with potential stem cell capacity represents an important step towards stable gene transfer into HSCs and gene therapy of blood disorders. Furthermore, the virological aspects of this invention contribute to a better understanding of adenovirus cell interactions.

**I: Retargeting of Ad5 based vectors with modified fibers carrying specific ligand peptides for HSC and other cell types**

5 Another alternative to make Ad5-capsid-based vectors suitable for HSC gene therapy is to incorporate the coding sequence for HSC specific peptides into the H1 loop region of the Ad5 fiber gene. The modification of the H1-loop was successfully exercised by Krasnykh et al. with a 7 amino-acid long FLAG peptide (DYDDDDK). Using phage display peptide libraries (Pascqualini, R. et al., 1996, *Nature*, 380:364-66), Renata  
10 Pasqualini (La Jolla Cancer Research Center) reported recently, at the First Meeting of the American Society for Gene Therapy, the identification of small peptide ligands specific for bone marrow cells. The corresponding sequences encoding these peptides can be added to modify the H1 loop sequence employing site-directed mutagenesis. Optimally, the ligands should allow for the efficient internalization of adenoviral particles  
15 based on a CAR- and integrin independent pathway. Modified adenoviral vectors containing the AAVBG cassette can be produced and tested for HSC tropism as described above.

**Adenovirus peptide display:**

20

In order to retarget adenoviruses to any cell type of interest, a strategy is provided which involves creating a library of adenoviruses displaying random peptides in their fiber knobs as ligands and screening this library for adenovirus variants with tropism to a particular cell type in vitro and potentially in vivo.

25

The development of the adenovirus peptide display technique is based on the following ideas. (i) Although the tertiary structure of the Ad5 fiber knob is known, it remains unclear which domains are involved in receptor binding. There are data suggesting that receptor-binding domains partially overlap with hemagglutination domains, which are  
30 well characterized for a number of serotypes. Therefore, three intramolecular loop regions representing potential receptor binding sites can be substituted by random peptide

libraries. Eight amino acid residues in the center of the FG, or GH loops can be substituted by octameric random peptides (Figs. 26 and 27). These substitutions will replace CAR tropism and allow for infection of refractory cell types. (ii) To synthesize the oligonucleotides encoding the peptide library a novel technique to assemble pre-synthesized trinucleotides representing the codons for all 20 amino acids is employed. This avoids termination codons and assures optimal codon usage and translation in human cells. Synthesis of a completely randomized library is possible with all 20 amino acids being incorporated with the same probability and a partially randomized library with only three (in average) random amino acids substitutions per octamer at random positions with a random amino acid to maintain certain critical features of the tertiary knob structure while introducing variability. The last model is based on the distribution of amino acids present in the hypervariable CDR 1 or 2 region of immunoglobulins. (iii) To maintain a representative library size of about  $10^{10}$  different octamers per modified loop, a new cloning strategy is employed to allow for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. This strategy is based on a "seamless" cloning technique available from Stratagen. (iv) In order to produce the library of viruses, viral genomic DNA containing the modified fiber sequences is transfected into 293 cells without reduction of the library size. This critical step is done by conjugating the viral library DNA to carrier Ad5-based adenovirus via poly-lysine to assure 100% transfection efficiency. This technique allows for coupling of  $\sim 1\mu\text{g}$  of plasmid DNA (or  $\sim 1 \times 10^{10}$  adenoviral genomes) to  $10^{10}$  viral particles which can be used to infect 293/cre cells at an MOI of 10-100. Importantly, the carrier adenoviral genome has the packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express cre recombinase (293/cre). This helper virus system is routinely used to produce so-called gutless adenoviruses. Therefore, the virus progeny represents library genomes packaged into capsids containing preferentially Ad5 fibers. This is important for the next infection step into 293 cells at a MOI of 1 to assure a homogeneous fiber population on the capsid where the fibers are encoded by the packaged genome.

#### **J. Production Of Adenovirus Vectors With Increased Tropism To Hepatocytes**

An example of a G-H loop substitution to target Ad5 to hepacytes was successful. Preliminary tests demonstrated that two evolutionarily conserved regions within the malaria circumsporozoite surface protein (CS) termed RI and RII+ mediate specific  
5 interaction with hepatocytes but not with other organs (including spleen, lung, heart and brain), nor with Kupffer cells, liver endothelial cells or with other regions of the hepatocyte membrane (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Shakibaei, M. and U. Frevert, 1996, *J. Exp. Med.*, 184:1699-711). These regions are conserved among different species including *Plasmodium berghei*, *P. cynomogli*, and *P. falciparum* that  
10 infect mouse, monkey and human hepatocytes, respectively (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). Peptides derived from RI (KLKQPG) or RII (EWSPCSVTCGNGIQVRIK) blocked CS binding to hepatocytes and infection by sporozoites in vivo ((Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). RI and RII+ peptides were  
15 separately inserted into Ad5-fiber knob (H-I and G-H loop) containing mutation with abolished binding to CAR and alpha-v integrins (Kirby, L. et al., 2000, *J. Virol.*, 74:2804-13; Wickham, T. J. et al., 1995, *Gene Ther.*, 2:750-6). Based on preliminary data, a short-shafted fiber was used so that the virus entry strategy predominantly depends on the interaction with the primary (hepatocyte-specific receptor). The hepatocyte-specific  
20 ligands are flanked by short glycine stretches to provide flexibility and embedded into a loop formed by two cystines. This is one of the classical strategies to incorporate ligands into a protein scaffold (Doi, N. and H. Yanagawa, 1998, *Cell Mol. Life Sci.*, 54:394-404; Koivunen, E. et al., 1995, *Biotechnology (NY)*, 13:265-70) and to guarantee their presentation at the protein surface. The biodistribution of the best variants is tested in  
25 vivo in C57B1/6 mice based on Southern blots or PCR for vector DNA in different organs. This mouse strain is known to be susceptible to infection with *P. berghei* (Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81).

**K. Production Of Adenovirus Vectors With Increased Tropism To Tumor Cells**

A similar strategy is to insert two peptides obtained after selection for tumor tropism by displaying random peptides on filamentous phages. The first double cyclic peptide  
5 (RGD-4) proved to bind specifically to integrins present on tumor vasculature (Ellerby H. M. et al., 1999, *Nat. Med.*, 5:1032-8). The second peptide targets specific matrix metalloproteinases associated with metastatic tumor cells as shown for the breast cancer cell line MDA-MB-435 (Koivunen, E. et al., 1999, *Nat. Biotechnol.*, 17:768-74). Tropism-modified vectors are tested in animal models with hepatic metastases derived  
10 from MDA-MB-435 cells (Fig. 28).

**L. Development Of A Peptide Display Technique Based On Adenoviruses**

A synthetic peptide library is described that allows adenovirus vectors to express random  
15 peptides in the G-H loop of the fiber knob domain. The technique of a phage display library is optimized to generate a library of adenoviruses displaying random peptides in their fiber knob. This library of adenovirus variants is then screened for tropism to a particular cell type in vitro and potentially in vivo. The oligonucleotides encoding the peptide library employ a novel technique to assemble pre-synthesized trinucleotides  
20 representing the codons for all 20 amino acids. This will end the termination codons and assure optimal codon usage and translation in human cells. To maintain a representative library size, a new "seamless" cloning strategy that allows for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. Transfection into 293 cells is done by  
25 conjugating the viral library DNA to carrier Ad5-based adenovirus via polylysine to assure a 100% transfection efficiency. Importantly, the carrier adenoviral genome has its packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express Cre recombinase (293/cre). The library is produced with E1-positive viruses depleted for CAR and integrin tropism. Only variant that have  
30 successfully infected the cell type of interest will replicate, resulting in de novo produced

virus. The sequence of the peptide ligand that conferred the particular tropism will then be analyzed in do novo produced virus.

### EXAMPLE III

5

#### COMBINATION NOVEL ADENOVIRAL VECTOR AND MODIFIED FIBER PROTEIN

This example describes the following studies which combine the technology of the integrating adenovirus vector that is devoid of all adenoviral genes with the modified fiber protein that retargets the vector to quiescent HSC.

##### A. Transduction studies with re-targeted vectors in HSC:

15 In order to transduce quiescent HSC and integrate into chromosomal DNA, retargeted  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors are tested for reporter gene expression, and vector integration simultaneously while analyzing their clonogenic capacity. The modified  $\Delta\text{Ad.AAV}^{\text{fx}}$  hybrid vectors contain genomes devoid of all adenoviral genes (a "gutless" adenovirus vector) packaged into Ad5 capsids with modified fibers. Rep may be incorporated into  
20 these  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors to allow for site-specific integration into AAVS1.

##### Transduction studies:

Purified human CD34+ cells in IMDM/FCS+IL-3 and SCF are infected with different  
25 doses of  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-BG}$  ( $1\text{-}10^7$  genomes per cell). CD34+ cells infected with  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-}\beta\text{Gal}$  are cultured for 2 days in suspension and sort  $\beta\text{-Gal}^+$  cells by FACS using FDG as substrate. This determines the infection efficiency.  $\beta\text{-gal}$  expressing cells are then submitted to clonogenic assays in semi-solid cultures (in two dishes per MOI) in the presence of multiple cytokines. (IL-3, SCF, Epo, G-CSF, GM-CSF, IL-7, Tpo). A  
30 first set of semi-solid cultures can be evaluated after 7 days; another set can be analyzed after 14 days. Colonies that have formed in semisolid culture can be characterized by



light microscopy and subsequently stained with X-Gal staining. Most of the vector genomes should remain episomal and can be lost with successive cell divisions. Thus, while most cells can be X-Gal positive at day 2 or day 7 after infection, most of the larger colonies (analyzed at day 14 p.i.) may not stain homogeneously for  $\beta$ -Gal. A representative number of X-Gal positive and X-Gal negative colonies can be picked and analyzed for episomal and integrated vector DNA. The outcome depends on the MOI used for infection and the integration status of the vector. These studies determine whether hybrid vectors can infect primitive progenitors.

10 Detailed characterization of hybrid vector integration:

CD34+ cells can be infected with  $\Delta$ Ad.AAV<sup>fx</sup>-SNori (MOI 1-10<sup>7</sup>) and subjected to G418 selection in methyl cellulose (MC) cultures in the presence of growth factors (IL-3 and SCF). The resulting colonies are a mixture of mainly myeloid cells. The number and morphology of G418 resistant colonies can be determined after 2 weeks of selection. This strategy may be disadvantageous in that the appropriate stem cell may not divide and form G418 resistant colonies under the specific culture condition used. Moreover, it may be difficult to perform G418 selection on a population of heterogenous cells, which vary in their sensitivity to G418. Therefore, another set of  $\Delta$ Ad.AAV<sup>fx</sup>-SNori infected CD34+ cells can be cultured in methyl-cellulose (+IL-3, SCF) without G418 selection. After 2-3 weeks, single colonies can be picked from both (w/ and w/o G418) MC cultures, morphologically characterized, and analyzed for integrated vector using the modified protocol developed for integration studies in a small number of cells (see Figure 8). This strategy allows the assessment of whether hybrid vectors integrate into the genome of CD34+ cells cultured in the presence of growth factors. This study gives us an idea about potential position effects affecting neo or  $\beta$ gal expression from integrated vector copies and about the structure of the integrated vector and the flanking chromosomal regions.

30 An alternative method to confirm vector integration:

Fluorescence *in situ* hybridization (FISH) analysis, can be performed in individual cells from MC colonies. CD34+ cells are cultured in MC in the presence of growth factors to induce cell division and subsequently treated with colchicine. Metaphase chromosome spreads are analyzed with biotin-ATP labeled probe specific for the  $\beta$ Gal or SNori gene and a dioxigenin-UTP labeled probe for the human X-chromosome as an internal control (provided by Christine Distech, University of Washington). Specific hybridization can be visualized with corresponding anti-biotin or anti-DIG antibodies labeled with different fluorochromes (e.g. FITC and Texas Red). Hybrid vector DNA may integrate as concatemers, which would facilitate detection by FISH. This technique allows one to localize the chromosomal integration sites of hybrid vectors.

Test transduction into quiescent bone marrow subpopulations:

Hybrid vectors described so far can be tested to see whether quiescent CD34+ cells can be stably transduced. To avoid significant cell proliferation, purified CD34+ cells are cultured in serum free IMDM supplemented with thrombopoietin (Tpo). Tpo can alone support the survival of stem cells without stimulating their active cell proliferation (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69). To analyze the proliferation status of CD34+ cells at the time point of infection with the hybrid vector  $\Delta$ Ad.AAV<sup>h</sup>BG, BrdU is added 2 hours before infection to the culture medium. One set of cells are maintained as suspension culture in IDAM containing Tpo only for two days. Another set of cells are grown in IDAM+Tpo supplemented with multiple cytokines. Forty eight hours after infection, CD34+ cells can be FACS sorted for beta Gal expression using FDG. FDG positive cells can be further analyzed for cellular DNA replication based on BrdU incorporation and for specific CD34+ subset markers. To do this, cytopins from FDG+ cells can be submitted to immunofluorescence with BrdU specific antibodies and with antibodies to specific cell surface markers (e.g. CD38, CD41). Alternatively, consecutive paraffin sections of the same cell can be analyzed for (a) transgene expression by X-Gal staining, (b) DNA synthesis based on BrdU incorporation, and (c) specific surface markers. This allows one to confirm that the culture conditions with Tpo alone prevent

significant genomic DNA replication and subsequent cell proliferation as well as to determine whether quiescent CD34+ cells can be infected based on beta Gal expression in cells where BrdU labeling is absent.

5 Test hybrid vectors integration into quiescent CD34+ cells:

Two sets of CD34+ cells are infected. The first set of  $\Delta$ Ad.AAV<sup>fix</sup>SNori infected cells are cultured for 5-7 days in the presence of cytokines; the other set is cultured without cytokines. To maintain CD34+ cell viability without cytokines during this period, the  
10 cells are cultured in the presence of Tpo or underlaid with a stromal cell line (AFT024) (Moore, K. et al., 1997, *Blood*, 89:4337-47), which can maintain HSC viable for 4 to 7 weeks. After this specific time period, both sets are submitted to clonogenic assays (in the presence of multiple cytokines) either in combination with G418 selection or without selection. Single colonies are analyzed morphologically and submitted to genomic DNA  
15 analysis (Figure 8) to determine the vector integration status. The ultimate proof for stem cell transduction is the in vivo survival/expansion assay. To do this, the CD34+ cells expressing beta Gal are used for transplantation experiments. If the number of FDG+ cells is not sufficient, total  $\Delta$ Ad.AAV<sup>fix</sup>BG infected cells as well as all  $\Delta$ Ad.AAV<sup>fix</sup>-SNori infected cells can be used directly without selection. Transplantation can be performed  
20 via tail vein injection into sublethally irradiated SCID NOD mice (Dao, M.A., et al., 1998, *Blood*, 4, 1243-1255; Matsunaga, T. et al., 1998, *Blood*, 92:452-61). At different time points after transplantation (4 to 8 weeks), mice can be sacrificed to obtain bone marrow cells which then can be cultured in suspension until various assays are performed for X-Gal and cell markers as described earlier. These cells also can be submitted to a  
25 secondary colony assay in MC or secondary transplantation into SCID NOD mice. Furthermore, MC colonies derived from these cells can be analyzed for the presence of integrated vector DNA by the method illustrated in Figure 8. The expression and integration data together allow conclusions about the repopulation efficiency and about potential position effects.

30

**B. Optimization of  $\Delta$ Ad.AAV<sup>fx</sup> vectors for  $\gamma$  r-globin expression in hematopoietic stem cells:**

One specific example of the invention is (a) to construct retargeted hybrid vector with the  $\gamma$ -globin as the transgene under the control of erythroid cell specific promoter, (b) to  
5 analyze the level and kinetics of  $\gamma$ -globin expression after transduction with hybrid vectors in *in vitro* and *in vivo* assays, (c) if required, to protect gene expression from position effects using  $\gamma$ -globin LCRs or insulators incorporated into hybrid vectors, and (d) to study whether  $\gamma$ -globin introns or heterologous introns can increase  $\gamma$ -globin  
10 expression.

Another central issue of the invention is to demonstrate that hybrid vectors can accommodate larger transgenes than rAAV and retroviruses. The insert size limitation of these vectors is 5kb. Transgene cassettes up to 8kb can be inserted into hybrid vectors as  
15 described. The maximal insert size may be about 14kb, if hybrid vectors are produced on the basis of E2a and/or E4 deleted rAd vectors in corresponding packaging cell lines. The maximal insert size in hybrid vectors is dictated by the packaging limit of first generation vectors (Ad.AAV) (<36kb) which are necessary intermediates for hybrid virus production at large scale. It is expected that stability and titer of Ad.AAV vectors with an 8kb globin  
20 gene cassette is comparable to the vector containing the 2.5-3.5kb cassette used in Ad.AAVBG, Ad.AAV1, and Ad.AAVSNori. The following example experiments address these issues.

**Production of  $\Delta$ Ad.AAV<sup>fx</sup> with large globin expression cassettes:**

25 In order to improve the condition of sickle cell disease, the expression level of the transferred  $\gamma$ -lobin gene must be at least 50% of that of each endogenous  $\beta$ gene. These levels of transgene expression can only be achieved by using optimal expression cassettes, including extended LCRs and intron containing gamma genes. (Forrester, W.C., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 1359-1363; Fraser, P., et al., 1998,  
30 *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in*

*Hematology*, 35, 105-111; Martin, D. et al., 1996, *Current Opinion in Genetics and Development*, 6:488-95), So far, most of the  $\gamma$ -globin expression cassettes are designed for retroviral and rAAV vectors, thus, less than 5kb and have to be devoid of internal splice sites or poly adenylation signals. With integrating vectors described herein, it is possible to go beyond this size limitation. This allows one to improve  $\gamma$ -globin expression in bone marrow cells in terms of an adequate expression level and long term persistence. For this purpose,  $\gamma$ -globin constructs developed by Li et al (Emery, D. W., et al. 1999 *Hum Gene Ther* 10:877-88; Li, Q., et al. 1999. *Blood* 93:2208-16) or by Ellis et al (Ellis, J., et al., 1996, *EMBO J.*, 15, 562-568; Ellis, J., et al., 1997, *Nucleic Acids Res.* 25, 1296-1302) is chosen.

(i) The first cassette contains a  $\gamma$ -globin expression unit used in retroviral vectors. This allows for a direct comparison between the two systems. This construct includes the beta promoter from -127 to the beta initiation codon, which is connected in frame with the gamma coding region. This beta promoter is combined with the 300bp HS40 derived from the human alpha globin locus, which acts as a strong enhancer for globin expression. The globin gene is the 1.1kb version with intron 1 and partially deleted intron 2. A second cassette is generated containing the HS40 beta promoter and gamma globin gene with the complete 3.3kb gamma globin gene.

(ii) The second construct contains the 6.5kb beta  $\mu$ LCR, which confer a dominant chromatin opening activity and an adequate level of gamma globin expression in transgenic mice. The LCR is linked to the short 1.1kb version of the gamma globin gene or the complete 3.3kb gamma gene.

(iii) Additional globin expression cassette can be generated which include insulators, MARs or SARs, as well as other elements that can improve transgene expression from integrated vectors or in transgenic animals, like introns derived from the HPRT or hGH genes (Chung, J.H., et al., 1997, *Proc. Natl. Acad. Sci. USA* 94, 575-580; Dunaway, M., et al, 1993, *Mol. Cell. Biol.*, 17, 182-189; Felsenfeld, G., et al., 1996,

*Proc. Natl. Acad. Sci. USA* 93, 93840-9388; Klehr, D., et al., 1991, *Biochemistry*, 30, 1264-1270).

Transduction studies with  $\Delta$ Ad.AAVfx-globin vectors:

5

Transduction studies with globin-hybrid vectors are performed as described earlier (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13). Transduced CD34+ cells are submitted to differentiation in colony assays or analyzed *in vivo* expansion assays in SCID-NOD mice. MC-colonies or bone marrow cells from experimental mice are  
10 analyzed for globin expression. Gamma-globin expression is measured using fluorescent anti-gamma-globin antibodies. RNAase protection studies can be performed to specifically quantitate gamma globin mRNA in comparison with  $\gamma$ -globin RNA. For these studies around  $10^4$ - $10^5$  cells are needed per test.

15 Position effects:

In the absence of the LCR, globin genes are subjected to strong position effects when they are transferred into cultured CD34+ cells or erythroleukemic lines (Fraser, P., et al., 1998, *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in*  
20 *Hematology*, 35, 105-111). Another concern is that site-specific integration of  $\Delta$ Ad.AAV/rep vectors into AAVS1 may silence transgene expression. If silencing happens, it can be overcome by incorporating LCRs such as the 6.5kb  $\gamma$  globin  $\mu$ LCR (Ellis, J., et al., 1996, *EMBO J.*, 15, 562-568; Grosveld, F., et al., 1998, *Seminars in*  
25 *Hematology*, 35, 105-111) or insulators into  $\Delta$ Ad.AAV based expression units. Insulators are DNA elements that protect an integrated reporter gene from chromosomal position effects or that block enhancer activated transcription from a downward promoter. Insulator elements are known for *Drosophila melanogaster* genes (Gypsy, suppressor of Hairy wing, scs, scs', Fab-7). for the chicken beta-globin gene (HS4) and for the T cell receptor (BEAD1; 14, 21.25). Specifically, the *Drosophila* gypsy or the beta globin  
30 insulator can be inserted as two copies flanking the globin expression cassette into hybrid vectors. The position effects can be examined in transduced MC-colonies based on the

analysis of integrated vector DNA (see Fig. 29) and gamma-globin mRNA quantification. Analogous studies can be performed on transduced human bone marrow cells obtained after transplantation of infected CD34+ cells into SCID-NOD mice.

5 Intron effects on gamma-globin expression:

A number of reports reveal that the deletion of globin introns, particularly the second intron of the beta and gamma genes, decrease globin mRNA stability and thus the expression level (Antoniou, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9). RNA viruses  
10 such as onco-retro, lenti- and foami viruses are problematic as vehicles for intron-containing transgenes. Because  $\Delta$ Ad.AAV is a DNA virus, it should package globin introns and LCRs if necessary without the decreased titers and rearrangements observed with retroviral vectors.

## APPENDIX I

### HUMAN AND ANIMAL ADENOVIRUSES AVAILABLE FROM AMERICAN TYPE CULTURE COLLECTION

- 1: Adenovirus Type 21 ATCC VR-1099 (NIAID V-221-002-014)  
2: SA18 (Simian adenovirus 18) ATCC VR-943 Classification  
10 3: SA17 (Simian adenovirus 17) ATCC VR-942 Classification  
4: Adenovirus Type 47 ATCC VR-1309 Classification: Adenov  
5: Adenovirus Type 44 ATCC VR-1306 Classification: Adenov  
6: Avian adenovirus Type 4 ATCC VR-829 Classification: Ad  
7: Avian adenovirus Type 5 ATCC VR-830 Classification: Ad  
15 8: Avian adenovirus Type 7 ATCC VR-832 Classification: Ad  
9: Avian adenovirus Type 8 ATCC VR-833 Classification: Ad  
10: Avian adenovirus Type 9 ATCC VR-834 Classification: Ad  
11: Avian adenovirus Type 10 ATCC VR-835 Classification: A  
12: Avian adenovirus Type 2 ATCC VR-827 Classification: Ad  
20 13: Adenovirus Type 45 ATCC VR-1307 Classification: Adenov  
14: Adenovirus Type 38 ATCC VR-988 Permit: PHS permit requ  
15: Adenovirus Type 46 ATCC VR-1308 Classification: Adenov  
16: Simian adenovirus ATCC VR-541 Classification: Adenovir  
17: SA7 (Simian adenovirus 16) ATCC VR-941 Classification:  
25 18: Frog adenovirus (FAV-1) ATCC VR-896 Classification: Ad  
19: Adenovirus Type 48 (candidate) ATCC VR-1406 Classifica  
20: Adenovirus Type 42 ATCC VR-1304 Classification: Adenov



- 21: Adenovirus type 49 (candidate) ATCC VR-1407 Classifica  
22: Adenovirus Type 43 ATCC VR-1305 Classification: Adenov  
23: Avian adenovirus Type 6 ATCC VR-831 Permit: USDA permi  
24: Avian adenovirus Type 3 (Inclusion body hepatitis virus)  
5 25: Bovine adenovirus Type 3 ATCC VR-639 Classification: A  
26: Bovine adenovirus Type 6 ATCC VR-642 Permit: USDA perm  
27: Canine adenovirus ATCC VR-800 Classification: Adenovir  
28: Bovine adenovirus Type 5 ATCC VR-641 Permit: USDA perm  
29: Adenovirus Type 36 ATCC VR-913 Classification: Adenovi  
10 30: Ovine adenovirus type 5 ATCC VR-1343 Classification: A  
31: Adenovirus Type 29 ATCC VR-272 Classification: Adenovi  
32: Swine adenovirus ATCC VR-359 Classification: Adenoviru  
33: Bovine adenovirus Type 4 ATCC VR-640 Permit: USDA perm  
34: Bovine adenovirus Type 8 ATCC VR-769 Permit: USDA perm  
15 35: Bovine adenovirus Type 7 ATCC VR-768 Permit: USDA perm  
36: Adeno-associated virus Type 2 (AAV-2H) ATCC VR-680 Cla  
37: Adenovirus Type 4 ATCC VR-4 Classification: Adenovirus  
38: Adeno-associated virus Type 3 (AAV-3H) ATCC VR-681 Cla  
39: Peromyscus adenovirus ATCC VR-528 Classification: Aden  
20 40: Adenovirus Type 15 ATCC VR-661 Classification: Adenovi  
41: Adenovirus Type 20 ATCC VR-662 Classification: Adenovi  
42: Chimpanzee adenovirus ATCC VR-593 Classification: Aden  
43: Adenovirus Type 31 ATCC VR-357 Classification: Adenovi  
44: Adenovirus Type 25 ATCC VR-223 Classification: Adenovi

25

- 45: Chimpanzee adenovirus ATCC VR-592 Classification: Aden  
46: Chimpanzee adenovirus ATCC VR-591 Classification: Aden  
47: Adenovirus Type 26 ATCC VR-224 Classification: Adenovi  
48: Adenovirus Type 19 ATCC VR-254 Classification: Adenovi  
5 49: Adenovirus Type 23 ATCC VR-258 Classification: Adenovi  
50: Adenovirus Type 28 ATCC VR-226 Classification: Adenovi  
51: Adenovirus Type 6 ATCC VR-6 Classification: Adenovirus  
52: Adenovirus Type 2 Antiserum: ATCC VR-1079 AS/Rab (NIA  
53: Adenovirus Type 6 ATCC VR-1083 (NIAID V-206-003-014)  
10 54: Ovine adenovirus type 6 ATCC VR-1340 Classification: A  
55: Adenovirus Type 3 ATCC VR-847 (Derived from NIAID V-20)  
56: Adenovirus Type 7 ATCC VR-7 Classification: Adenovirus  
57: Adenovirus Type 39 ATCC VR-932 Classification: Adenovi  
58: Adenovirus Type 3 ATCC VR-3 Classification: Adenovirus  
15 59: Bovine adenovirus Type 1 ATCC VR-313 Classification: A  
60: Adenovirus Type 14 ATCC VR-15 Classification: Adenovir  
61: Adenovirus Type 1 ATCC VR-1078 (NIAID V-201-001-014)  
62: Adenovirus Type 21 ATCC VR-256 Classification: Adenovi  
63: Adenovirus Type 18 ATCC VR-1095 (NIAID V-218-003-014)  
20 64: Baboon adenovirus ATCC VR-275 Classification: Adenovir  
65: Adenovirus Type 10 ATCC VR-11 Classification: Adenovir  
66: Adenovirus Type 33 ATCC VR-626 Classification: Adenovi

- 67: Adenovirus Type 34 ATCC VR-716 Classification: Adenovi  
68: Adenovirus Type 15 ATCC VR-16 Classification: Adenovir  
69: Adenovirus Type 22 ATCC VR-257 Classification: Adenovi  
70: Adenovirus Type 24 ATCC VR-259 Classification: Adenovi  
5 71: Adenovirus Type 17 ATCC VR-1094 (NIAID V-217-003-014)  
72: Adenovirus Type 4 ATCC VR-1081 (NIAID V-204-003-014)  
73: Adenovirus Type 16 ATCC VR-17 Classification: Adenovir  
74: Adenovirus Type 17 ATCC VR-18 Classification: Adenovir  
75: Adenovirus Type 16 ATCC VR-1093 (NIAID V-216-003-014)  
10 76: Bovine adenovirus Type 2 ATCC VR-314 Classification: A  
77: SV-30 ATCC VR-203 Classification: Adenovirus, Simian (  
78: Adenovirus Type 32 ATCC VR-625 Classification: Adenovi  
79: Adenovirus Type 20 ATCC VR-255 Classification: Adenovi  
80: Adenovirus Type 13 ATCC VR-14 Classification: Adenovir  
15 81: Adenovirus Type 14 ATCC VR-1091 (NIAID V-214-001-014)  
82: Adenovirus Type 18 ATCC VR-19 Classification: Adenovir  
83: SV-39 ATCC VR-353 Classification: Adenovirus, Simian (  
84: Adenovirus Type 11 ATCC VR-849 (Derived from NIAID V-2  
85: Duck adenovirus (Egg drop syndrome) ATCC VR-921 Permi  
20 86: Adenovirus Type 1 ATCC VR-1 Classification: Adenovirus  
87: Chimpanzee adenovirus ATCC VR-594 Classification: Aden  
88: Adenovirus Type 15 ATCC VR-1092 (NIAID V-215-003-014)

- 89: Adenovirus Type 13 ATCC VR-1090 (NIAID V-213-003-014)  
90: Adenovirus Type 8 ATCC VR-1368 (Derived from NIAID V-20  
91: SV-31 ATCC VR-204 Classification: Adenovirus, Simian (  
92: Adenovirus Type 9 ATCC VR-1086 (NIAID V-209-003-014)  
5 93: Mouse adenovirus ATCC VR-550 Classification: Adenoviru  
94: Adenovirus Type 9 ATCC VR-10 Classification: Adenoviru  
95: Adenovirus Type 41 ATCC VR-930 Classification: Adenovi  
96: CI ATCC VR-20 Classification: Adenovirus, Simian (Mast  
97: Adenovirus Type 40 ATCC VR-931 Classification: Adenovi  
10 98: Adenovirus Type 37 ATCC VR-929 Classification: Adenovi  
99: Marble spleen disease virus (Hemorrhagic enteritis virus  
100: Adenovirus Type 35 ATCC VR-718 Classification: Adenovi  
101: SV-32 (M3) ATCC VR-205 Classification: Adenovirus, Sim  
102: Adenovirus Type 28 ATCC VR-1106 (NIAID V-228-003-014)  
15 103: Adenovirus Type 10 ATCC VR-1087 (NIAID V-210-003-014)  
104: Adenovirus Type 20 ATCC VR-1097 (NIAID V-220-003-014)  
105: Adenovirus Type 21 ATCC VR-1098 (NIAID V-221-011-014)  
106: Adenovirus Type 25 ATCC VR-1103 (NIAID V-225-003-014)  
107: Adenovirus Type 26 ATCC VR-1104 (NIAID V-226-003-014)  
20 108: Adenovirus Type 31 ATCC VR-1109 (NIAID V-231-001-014)  
109: Adenovirus Type 19 ATCC VR-1096 (NIAID V-219-002-014)  
110: SV-36 ATCC VR-208 Classification: Adenovirus, Simian (

- 111: SV-38 ATCC VR-355 Classification: Adenovirus, Simian (  
112: SV-25 (M8) ATCC VR-201 Classification: Adenovirus, Sim  
113: SV-15 (M4) ATCC VR-197 Classification: Adenovirus, Sim  
114: Adenovirus Type 22 ATCC VR-1100 (NIAID V-222-003-014)  
5 115: SV-23 (M2) ATCC VR-200 Classification: Adenovirus, Sim  
116: Adenovirus Type 11 ATCC VR-12 Classification: Adenovir  
117: Adenovirus Type 24 ATCC VR-1102 (NIAID V-224-003-014)  
118: Avian adenovirus Type 1 (Chicken Embryo Lethal Orphan: C  
119: SV-11 (M5) ATCC VR-196 Classification: Adenovirus, Sim  
10 120: Adenovirus Type 5 ATCC VR-5 Classification: Adenovirus  
121: Adenovirus Type 23 ATCC VR-1101 (NIAID V-223-003-014)  
122: SV-27 (M9) ATCC VR-202 Classification: Adenovirus, Sim  
123: Avian adenovirus Type 2 (GAL) ATCC V-280 Classificati  
124: SV-1 (M1) ATCC VR-195 Classification: Adenovirus, Simi  
15 125: SV-17 (M6) ATCC VR-198 Classification: Adenovirus, Sim  
126: Adenovirus Type 29 ATCC VR-1107 (NIAID V-229-003-014)  
127: Adenovirus Type 2 ATCC VR-846 Classification: Adenovir  
128: SV-34 ATCC VR-207 Classification: Adenovirus, Simian (  
129: SV-20 (M7) ATCC VR-199 Classification: Adenovirus, Sim  
20 130: SV-37 ATCC VR-209 Classification: Adenovirus, Simian (  
131: SV-33 (M10) ATCC VR-206 Classification: Adenovirus, Si

- 132: Avian adeno-associated virus ATCC VR-865 Classificatio  
133: Adeno-associated (satellite) virus Type 4 ATCC VR-646  
134: Adenovirus Type 30 ATCC VR-273 Classification: Adenovi  
135: Adeno-associated (satellite) virus Type 1 ATCC VR-645  
5 136: Infectious canine hepatitis (Rubarth's disease, Fox ence  
137: Adenovirus Type 27 ATCC VR-1105 (NIAID V-227-003-014)  
138: Adenovirus Type 12 ATCC VR-863 (Derived from NIAID V-2  
139: Adeno-associated virus Type 2 (molecularly cloned) ATCC  
140: Adenovirus Type 7a ATCC VR-848 (Derived from NIAID V-2

10

What is claimed:

1. A first generation recombinant adenovirus vector a portion of which integrates into a host cell genome, comprising:
  - 5 a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
  - d) At least one adenoviral sequence which directs adenoviral replication; and
  - 10 e) A right adenoviral inverted terminal repeat sequence,wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.
2. The adenovirus vector of claim 1, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype.
3. The adenovirus vector of claim 1, wherein the sequence which directs adenoviral replication comprises a sequence on the anti-parallel strand which encodes an adenoviral fiber protein including a fiber tail, a fiber shaft, and a fiber knob, wherein the fiber knob includes a G-H loop region.
4. The adenovirus vector of claim 3, wherein the sequence on the anti-sense strand which encodes the fiber tail is from the same serotype as the adenoviral inverted repeat sequence.
5. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:
  - 30 a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;

- c) A transgene sequence 3' to the polyadenylation sequence;
- d) A promoter sequence 3' to the polyadenylation sequence; and
- e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.

5

6. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:

- a) A left cassette inverted terminal repeat sequence;
- b) A promoter sequence 3' to the left inverted terminal repeat sequence;
- 10 c) A transgene sequence 3' to the promoter sequence;
- d) A polyadenylation sequence 3' to the transgene sequence; and
- e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.

- 15 7. The adenovirus vector of claim 5 or 6, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.

- 20 8. The adenovirus vector of claim 5 or 6, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.

9. The adenovirus vector of claim 8, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.

25

10. The adenovirus vector of claim 8, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.

11. The adenovirus vector of claim 8, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.

30



12. The transgene cassette of claim 5 or 6 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
- 5
13. The transgene cassette of claim 5 or 6, further comprising an insulator element.
14. The transgene cassette of claim 5 or 6, further comprising a bacterial origin of replication.
- 10
15. The adenoviral vector of claim 1, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
- 15
16. A first generation recombinant adenovirus vector which targets a host cell of interest and a portion of which integrates into the host cell genome so targeted, comprising two DNA strands, each strand being antiparallel to the other, the first strand comprising:
- 20
- a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
  - d) At least one adenoviral sequence which directs adenoviral replication; and
  - e) A right adenoviral inverted terminal repeat sequence,
- 25
- wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome, and wherein the second strand comprises a sequence which encodes an adenoviral fiber protein that permits targeting of the vector into the host cell of interest.

17. The adenovirus vector of claim 16, wherein the adenoviral protein includes a fiber tail, a fiber shaft, and a fiber knob, wherein the fiber knob includes a G-H loop region.
- 5 18. The adenovirus vector of claim 16, wherein the left and right adenovirus inverted terminal repeat sequences and the packaging sequence are from the same adenoviral serotype.
- 10 19. The adenovirus vector of claim 17, wherein the fiber tail is from the same serotype as the left and right adenoviral inverted repeat sequences.
20. The adenovirus vector of claim 17, wherein the fiber shaft is from a different serotype as the left and right adenoviral inverted repeat sequences.
- 15 21. The adenovirus vector of claim 20, wherein the fiber shaft is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
22. The adenovirus vector of claim 17, wherein the fiber shaft comprises a shortened length.
- 20 23. The adenovirus vector of claim 17, wherein the fiber knob is from a different serotype as the left and right adenoviral inverted repeat sequences.
24. The adenovirus vector of claim 23, wherein the fiber knob is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
- 25 25. The adenovirus vector of claim 17, wherein the fiber knob is a modified fiber knob protein comprising the G-H loop replaced with a heterologous peptide ligand sequence which binds to at least one surface protein on the host cell of interest.
- 30

26. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 5 b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - 10 e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
27. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 15 b) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the promoter sequence;
  - d) A polyadenylation sequence 3' to the transgene sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
  - 20
28. The adenovirus vector of claim 26 or 27, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 25 29. The adenovirus vector of claim 26 or 27, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.
- 30 30. The adenovirus vector of claim 24, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.

31. The adenovirus vector of claim 24, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.
32. The adenovirus vector of claim 24, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
33. The transgene cassette of claim 26 or 27 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
34. The transgene cassette of claim 26 or 27, further comprising an insulator element.
35. The transgene cassette of claim 26 or 27, further comprising a bacterial origin of replication.
36. The adenoviral vector of claim 16, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
37. A recombinant gutless adenovirus vector a portion of which integrates into a host cell genome, comprising:
- a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence; and
  - d) A right adenoviral inverted terminal repeat sequence,
- wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.

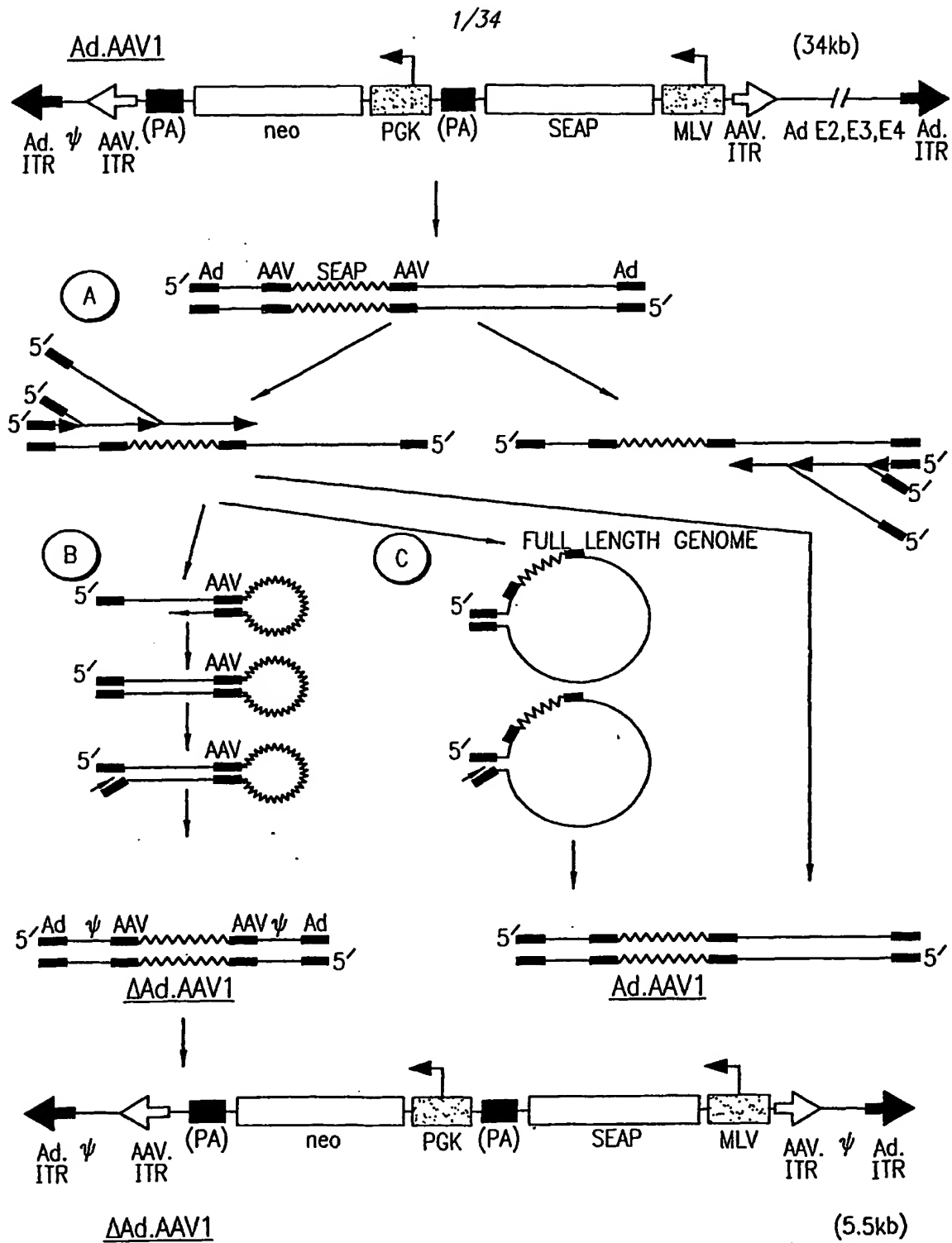
38. The adenovirus vector of claim 36, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype. (*this is the base vector*)
- 5 39. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - 10 c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 15 40. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- b) A left cassette inverted terminal repeat sequence;
  - c) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - d) A transgene sequence 3' to the promoter sequence;
  - 20 e) A polyadenylation sequence 3' to the transgene sequence; and
  - f) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 25 41. The adenovirus vector of claim 38 or 39, wherein the left and the right inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 30 42. The adenovirus vector of claim 38 or 39, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.

43. The adenovirus vector of claim 41, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.
44. The adenovirus vector of claim 41, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentomycin.
45. The adenovirus vector of claim 41, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
46. The adenovirus vector of claim 38 or 39 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
47. The adenovirus vector of claim 38 or 39, further comprising an insulator element.
48. The adenovirus vector of claim 38 or 39, further comprising a bacterial origin of replication.
49. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 5' portion of a gene of interest.
50. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 3' portion of a gene of interest.
51. A method of producing a resolved gutless adenovirus vector in a suitable cell, said method comprising introducing a first and a second adenovirus vectors of claim 1 or 16 into the cell under suitable conditions so that the recombinant adenovirus vectors undergo homologous recombination thereby producing a resolved gutless adenovirus vector.

52. A resolved gutless adenovirus vector produced by the method of claim 50.
53. The method of claim 50, wherein the first adenovirus vector comprises a transgene cassette having a 5' portion of a gene of interest, and wherein the  
5 second adenovirus vector comprises a transgene cassette having a 3' portion of the gene of interest, and wherein a part of the 5' portion overlaps with a part of the 3' portion so that homologous recombination occurs.
54. A method of producing a resolved gutless recombinant Ad vector by homologous  
10 recombination in a suitable cell, said method comprising contacting two parental recombinant Ad vectors, each comprising a transgene cassette containing a portion of a selected transgene with a region of overlapping homology, so that the first and second parental recombinant Ad vectors undergo homologous recombination at the region of overlapping homology, resulting in a resolved  
15 recombinant gutless Ad vector having both portions of the selected transgene, and wherein the selected transgene is within a transgene cassette flanked by a pair of ITRs.
55. A resolved gutless adenovirus vector produced by the method of claim 53.
- 20 56. An adenovirus library comprising a plurality of adenovirus vectors expressing fiber proteins which are displayed and modified by random peptide insertions.
57. The library of claim 55, wherein said fiber protein so displayed comprises a  
25 random peptide substituted in the G-H loop of the fiber protein knob domain.
58. A screening method for targeting adenovirus vectors for gene therapy comprising contacting the adenovirus library of claim 55 with a plurality of cells so that the cells are transduced with the adenovirus vectors of the adenovirus library  
30 transduction occurs and detecting the cells so transduced.

59. The adenovirus vector of claim 5, 6, 26, 27, 38, or 39 further comprising a nucleotide sequence encoding a rep78 protein.

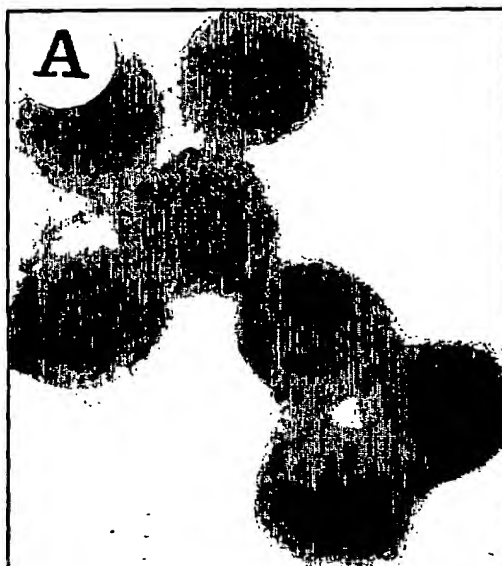




**FIG. 1**  
SUBSTITUTE SHEET (RULE 26)

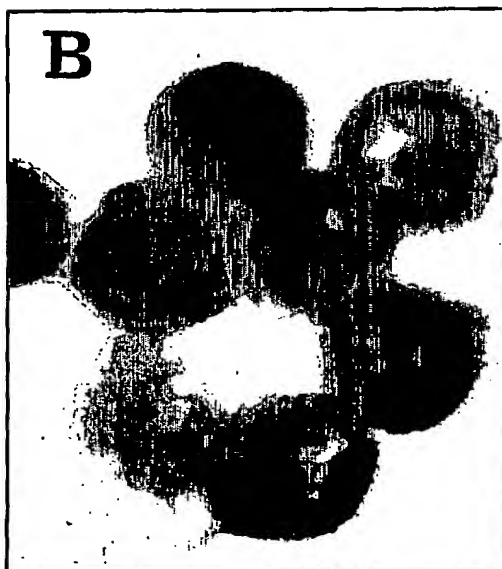
2/34

Ad.AAV  
(33kb)



50nm

$\Delta$ Ad.AAV  
(5.5kb)



50nm

*FIG. 2*

SUBSTITUTE SHEET (RULE 26)

3/34

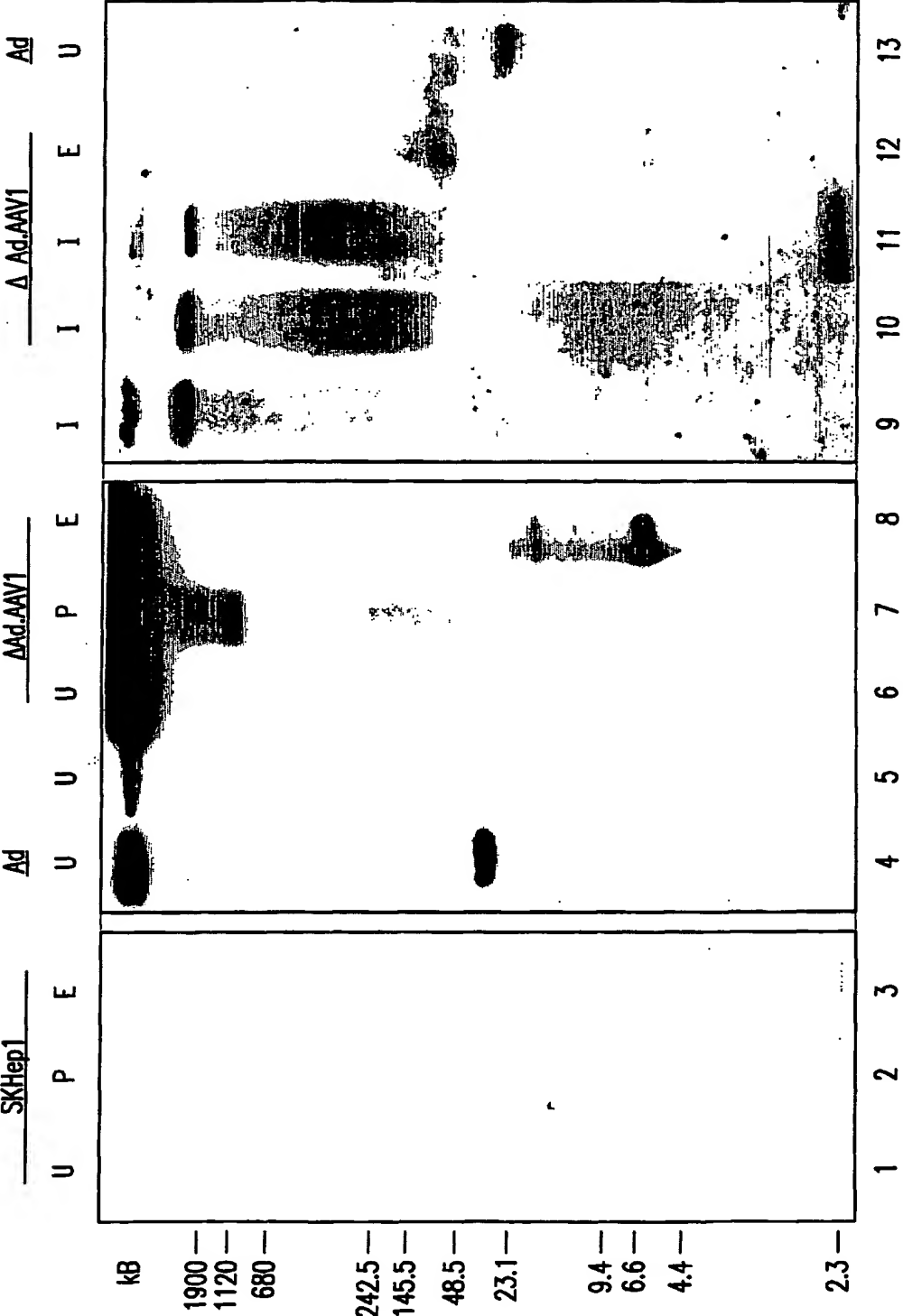
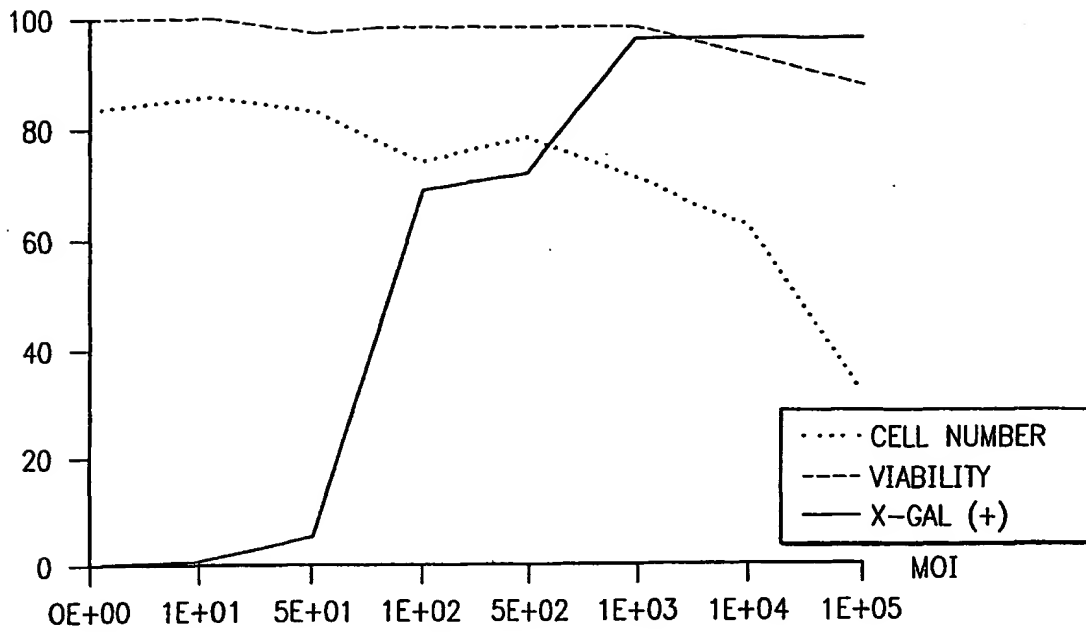
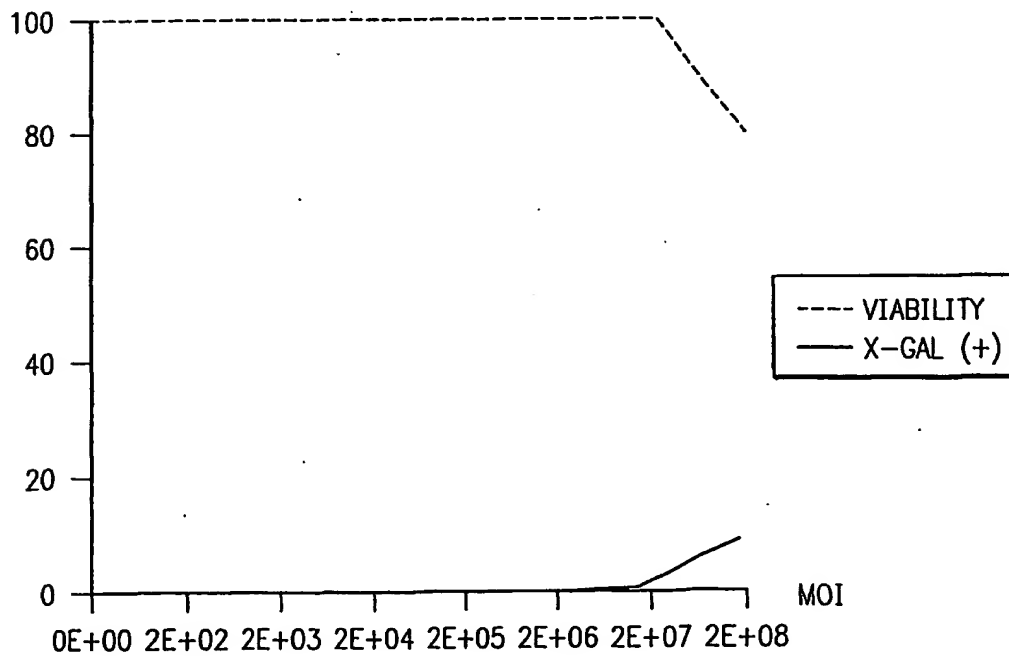


FIG. 3

4/34

FIG. 4A K562FIG. 4B CD34+

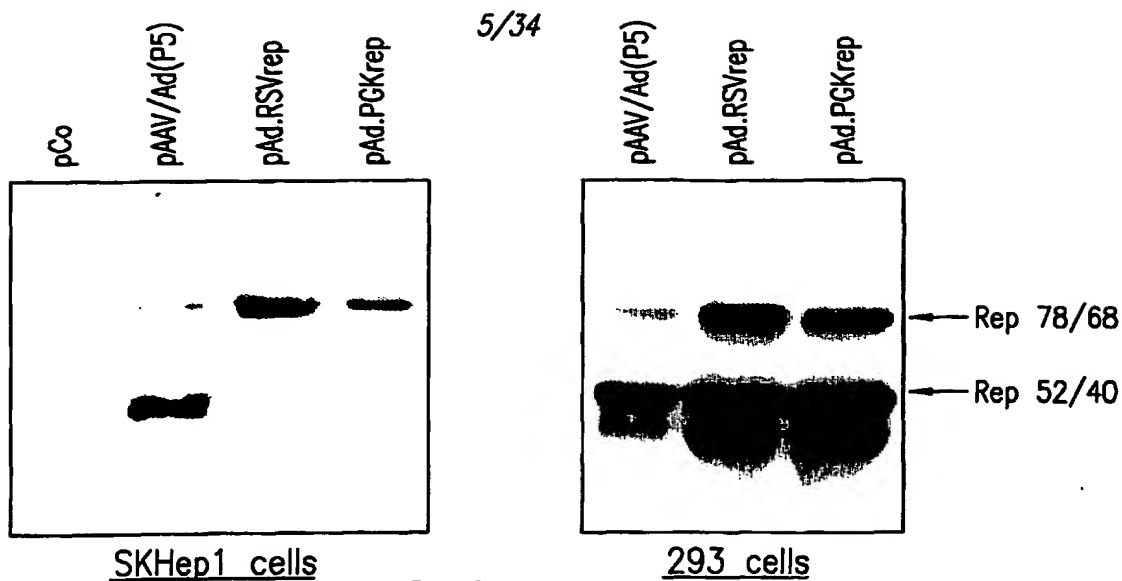


FIG. 5

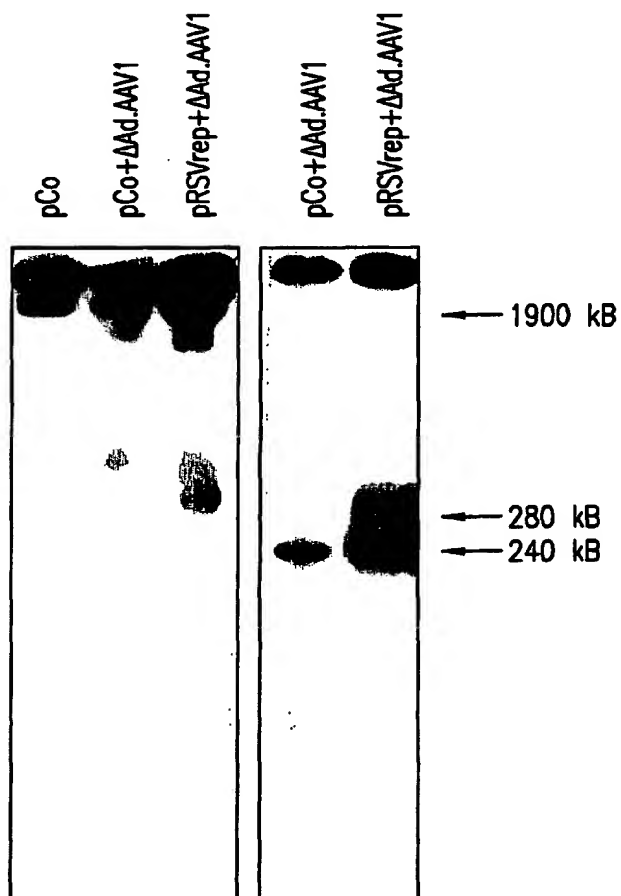


FIG. 6

SEAP probe AAVS1 probe

SUBSTITUTE SHEET (RULE 26)

6/34

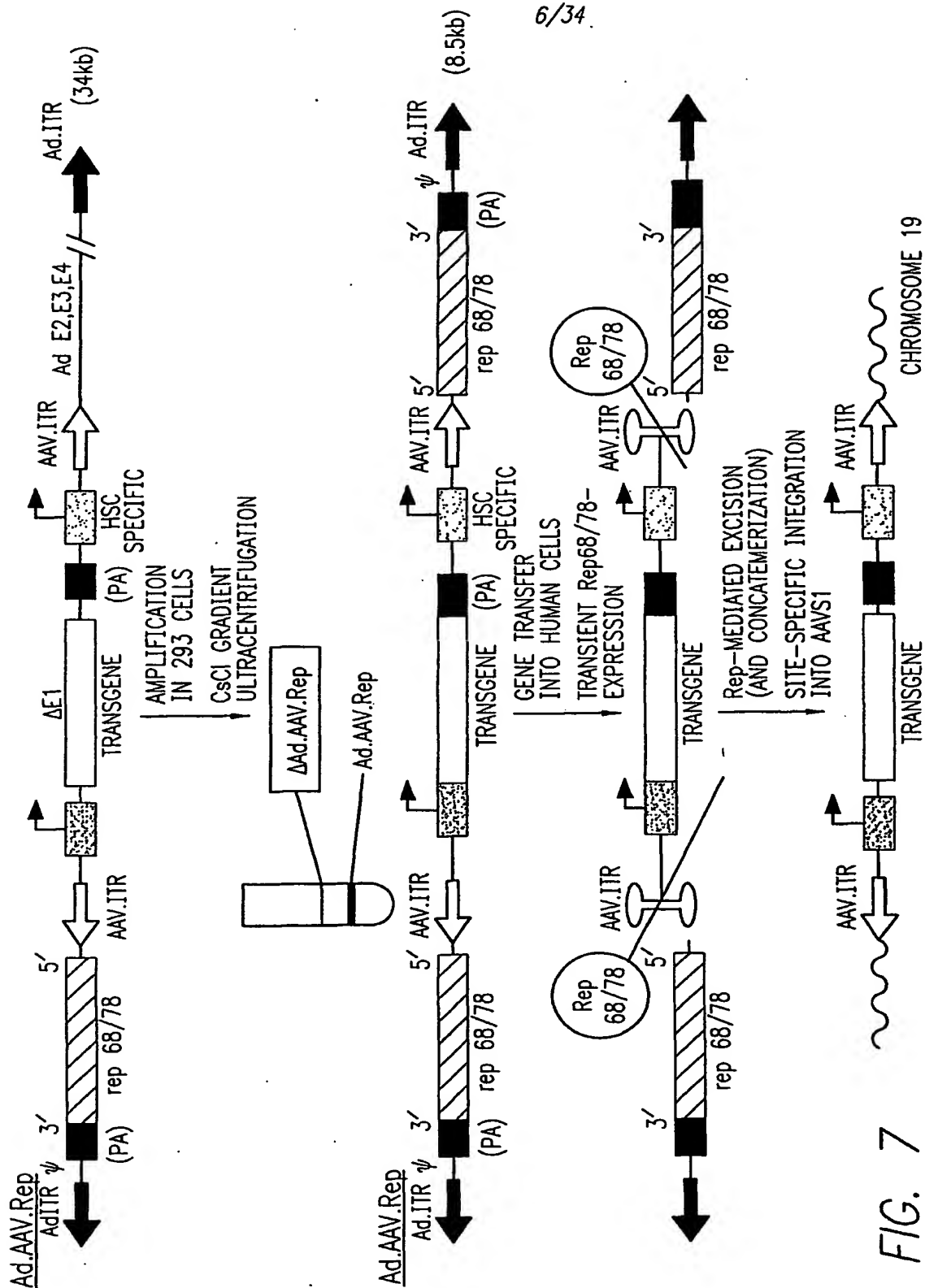


FIG. 7

7/34

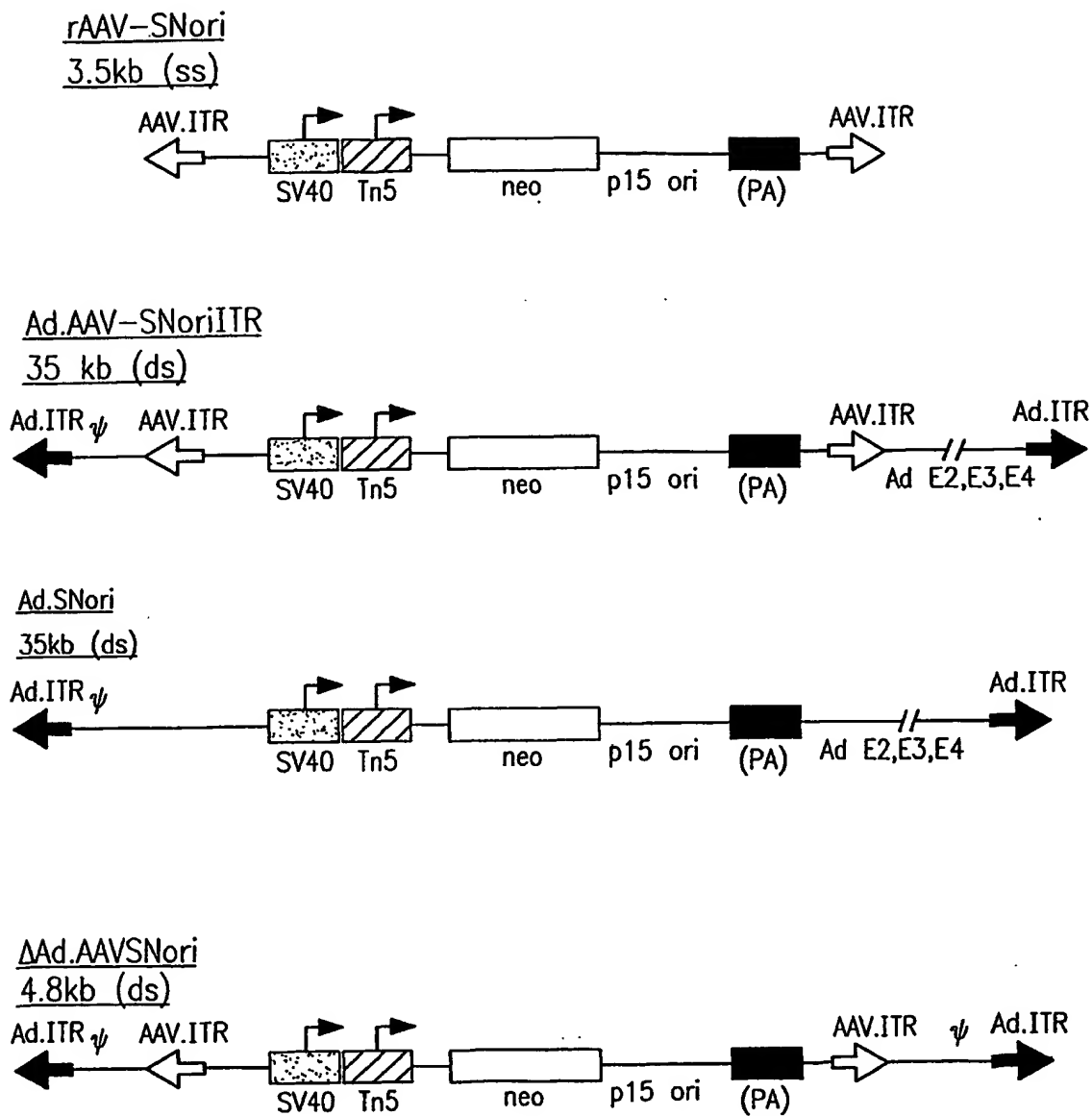


FIG. 8A

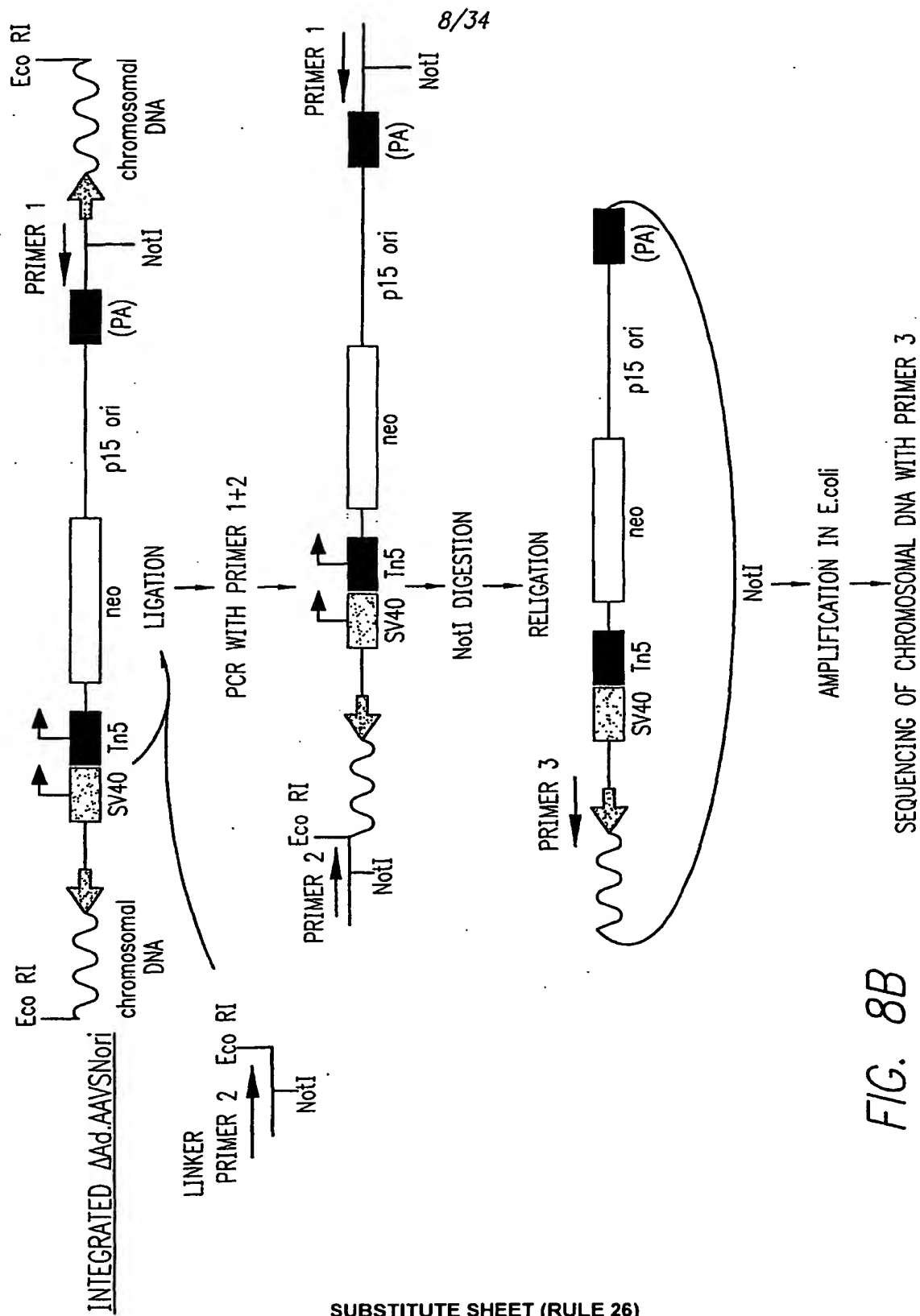


FIG. 8B



9/34

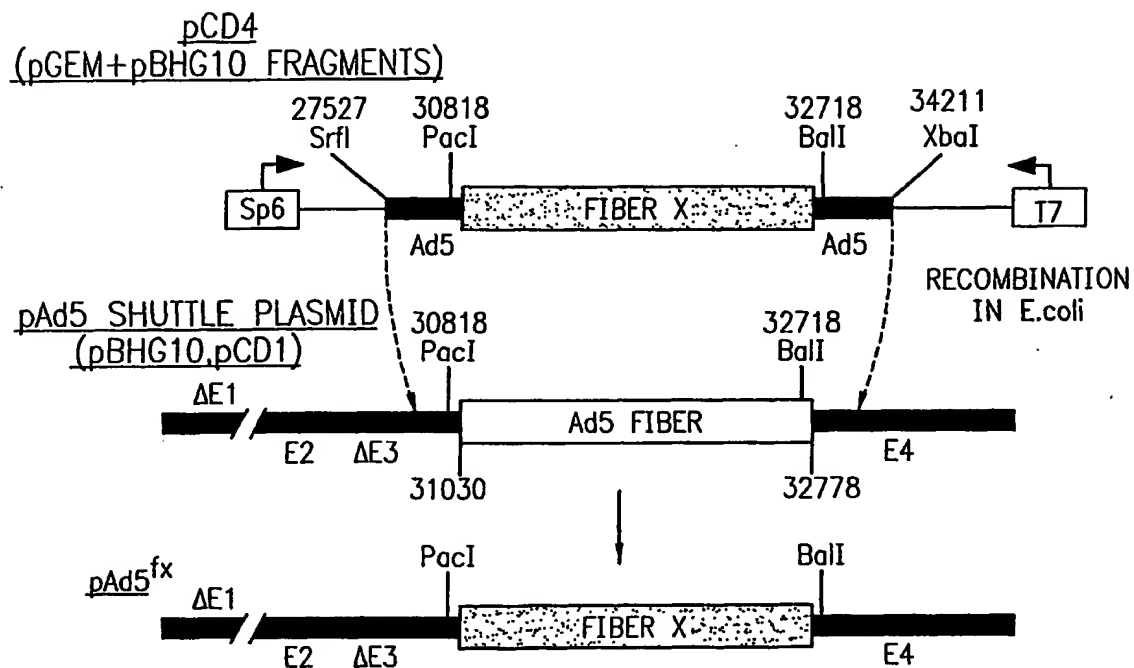


FIG. 9

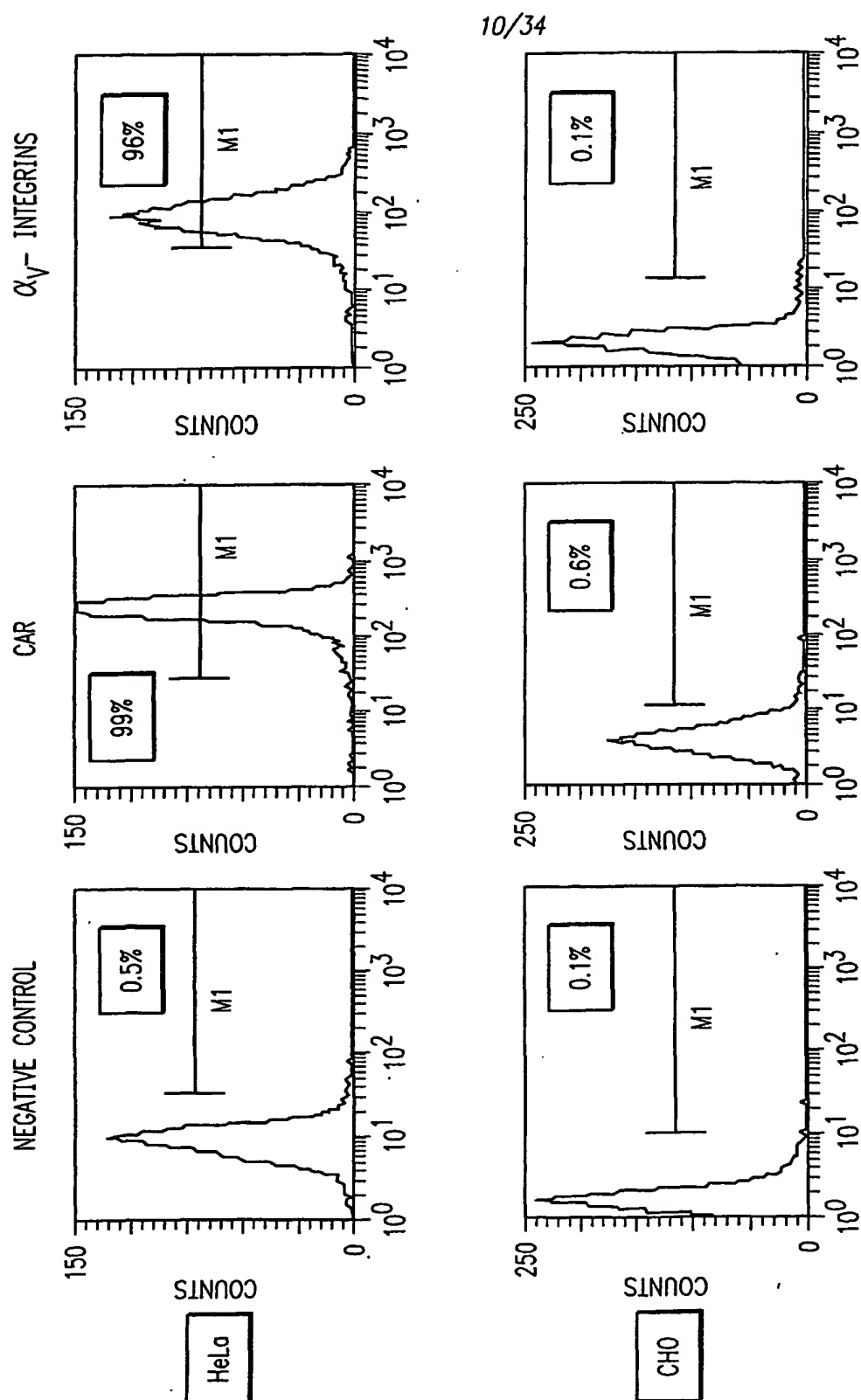


FIG. 10-i

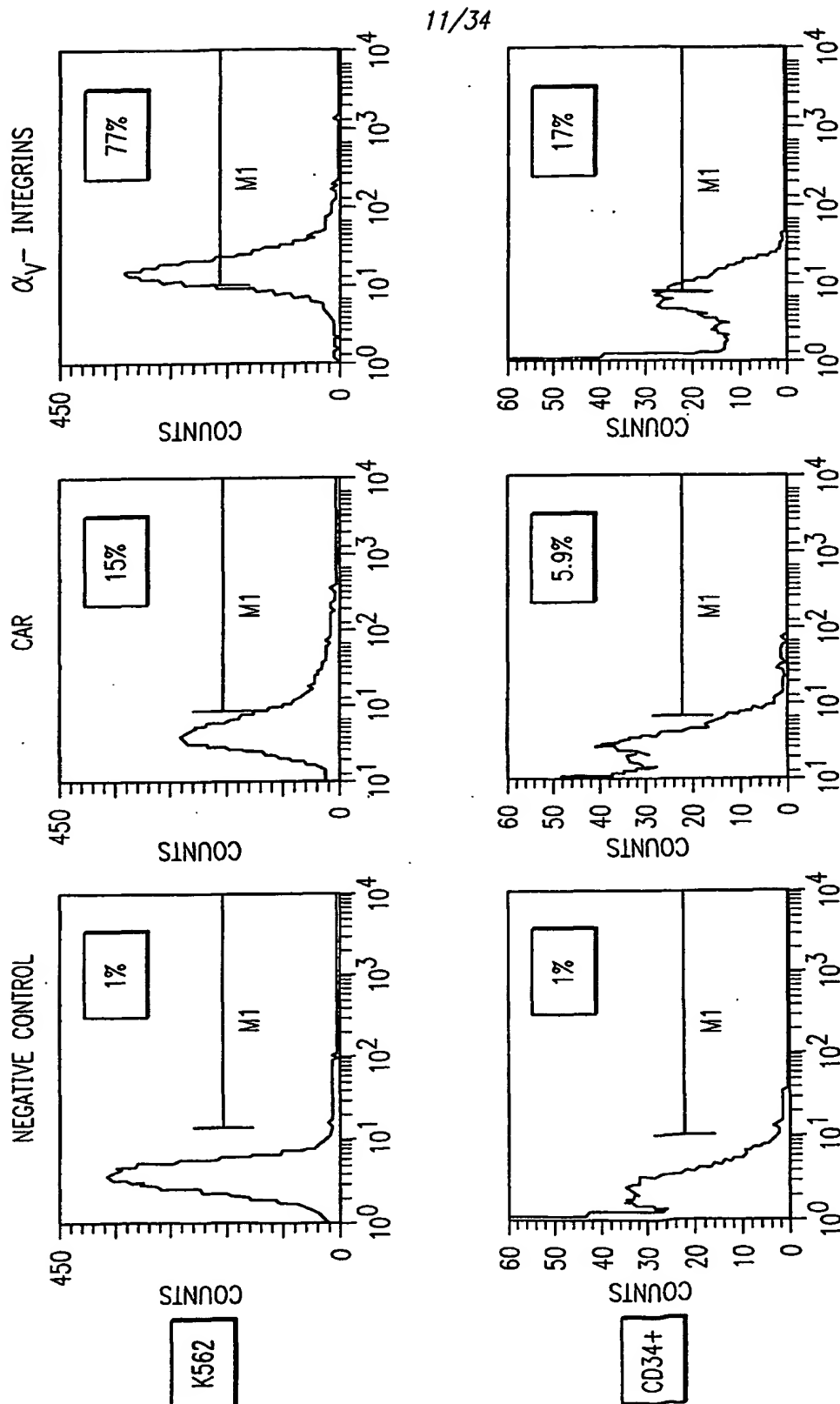
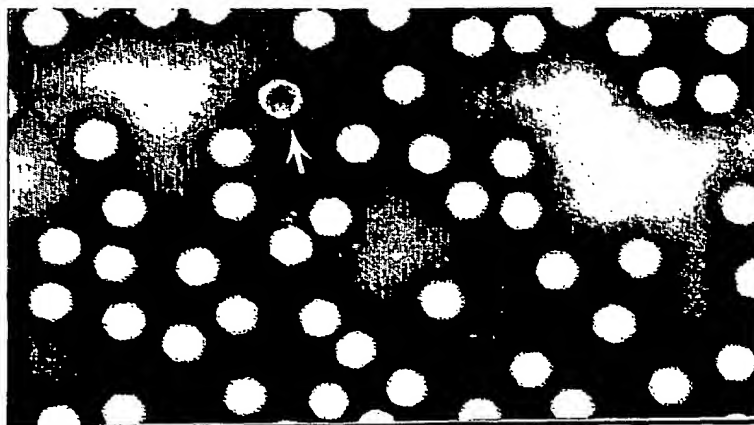
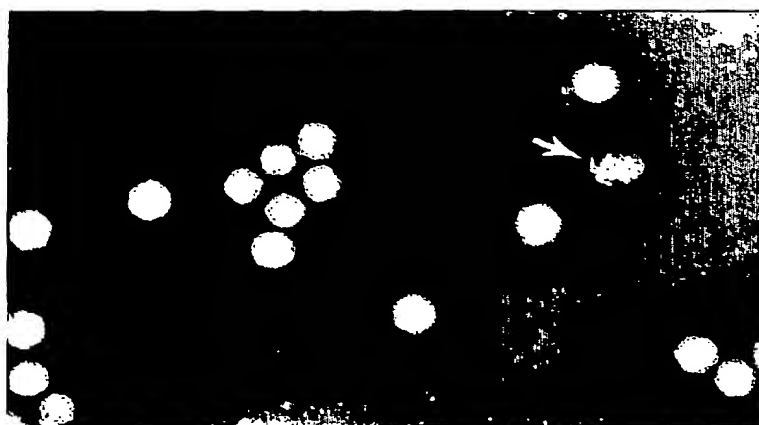


FIG. 10-ii

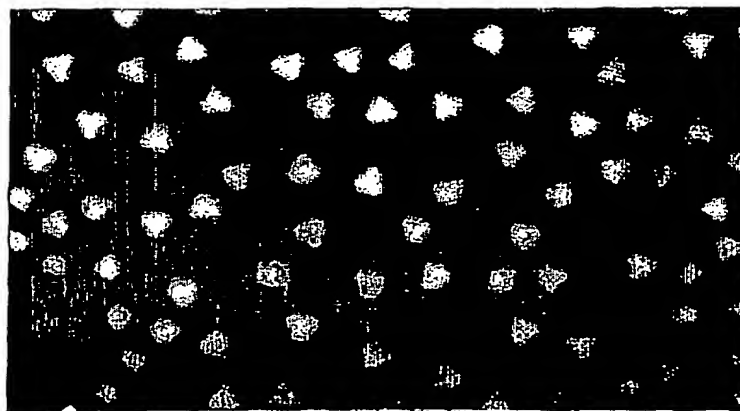
12/34



Ad 35



Ad 9



Ad 5

FIG. 11

13/34

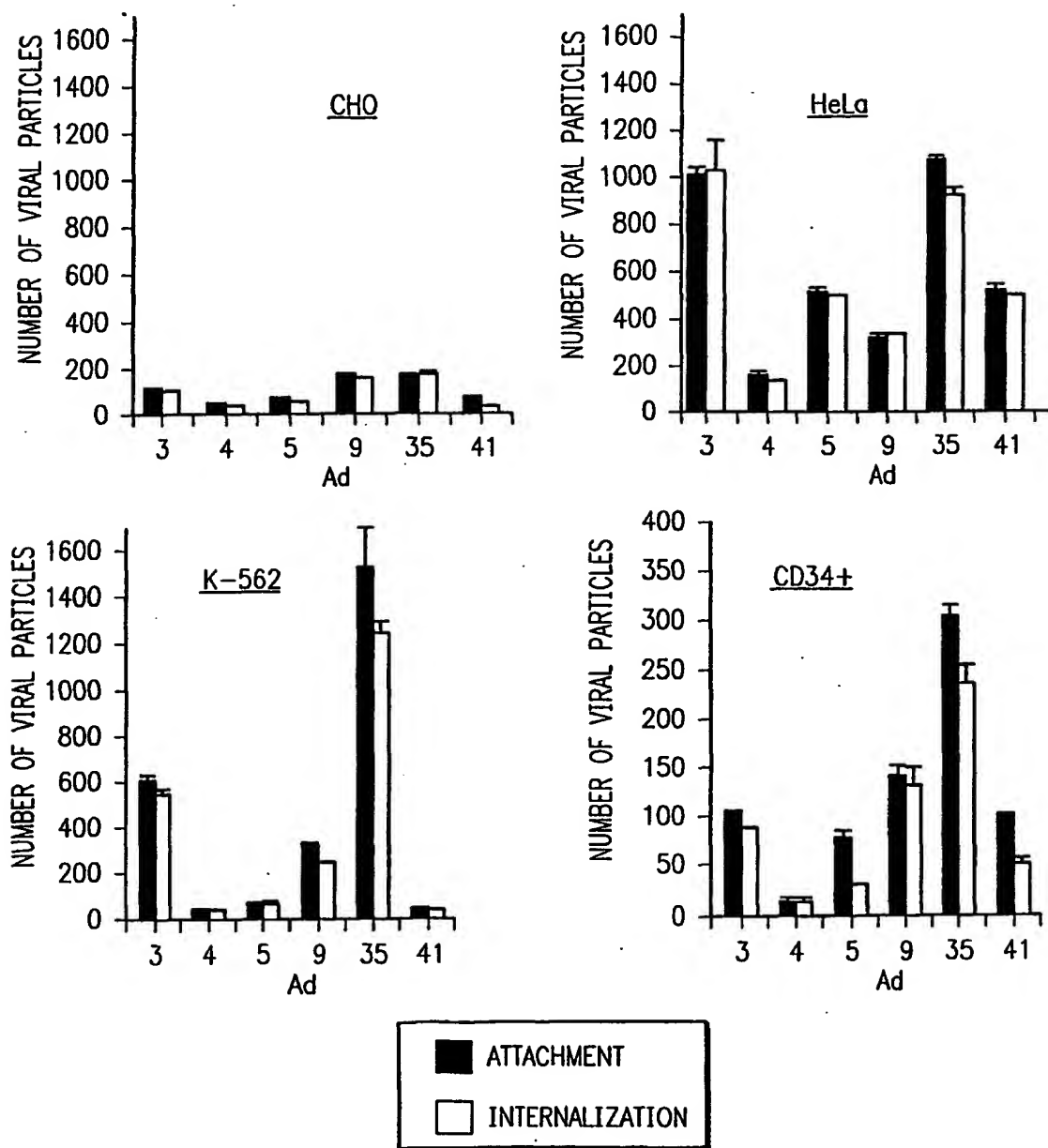
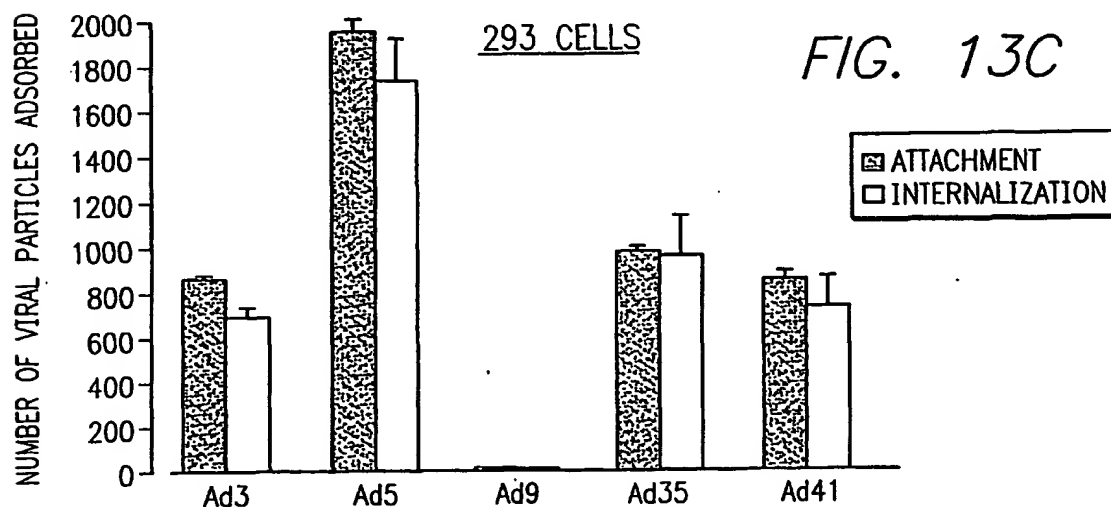
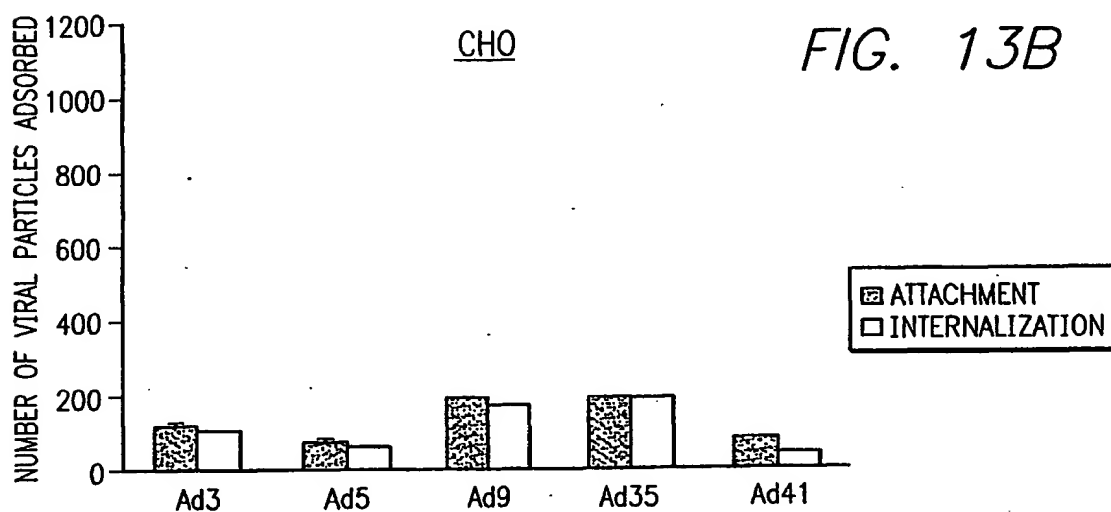
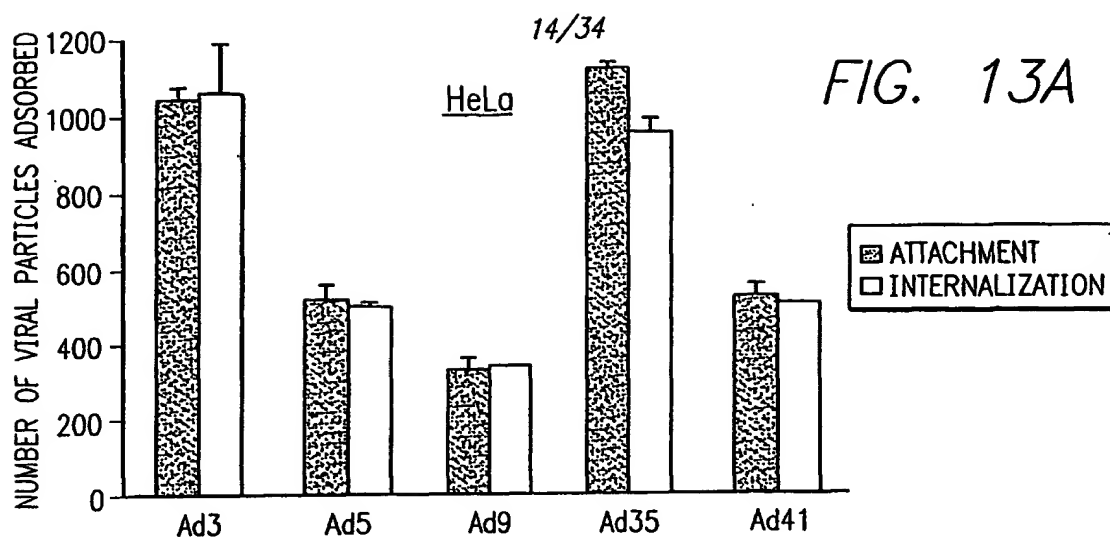


FIG. 12



15/34

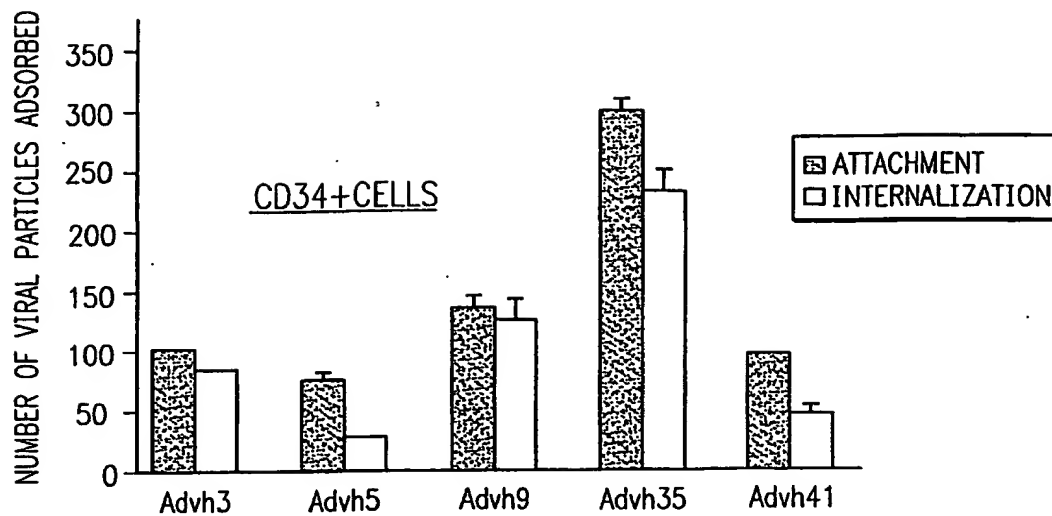


FIG. 14A

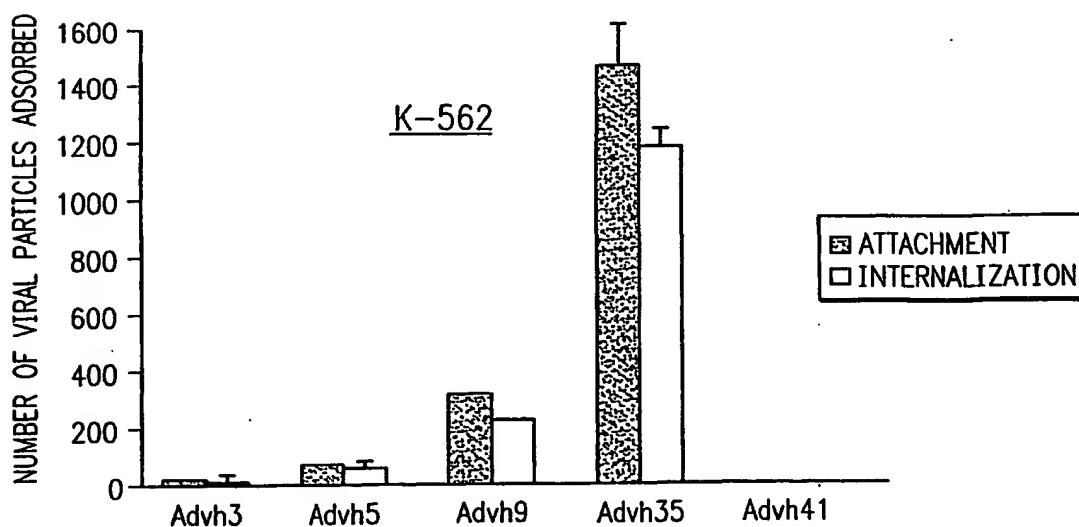


FIG. 14B

16/34

K562 cells

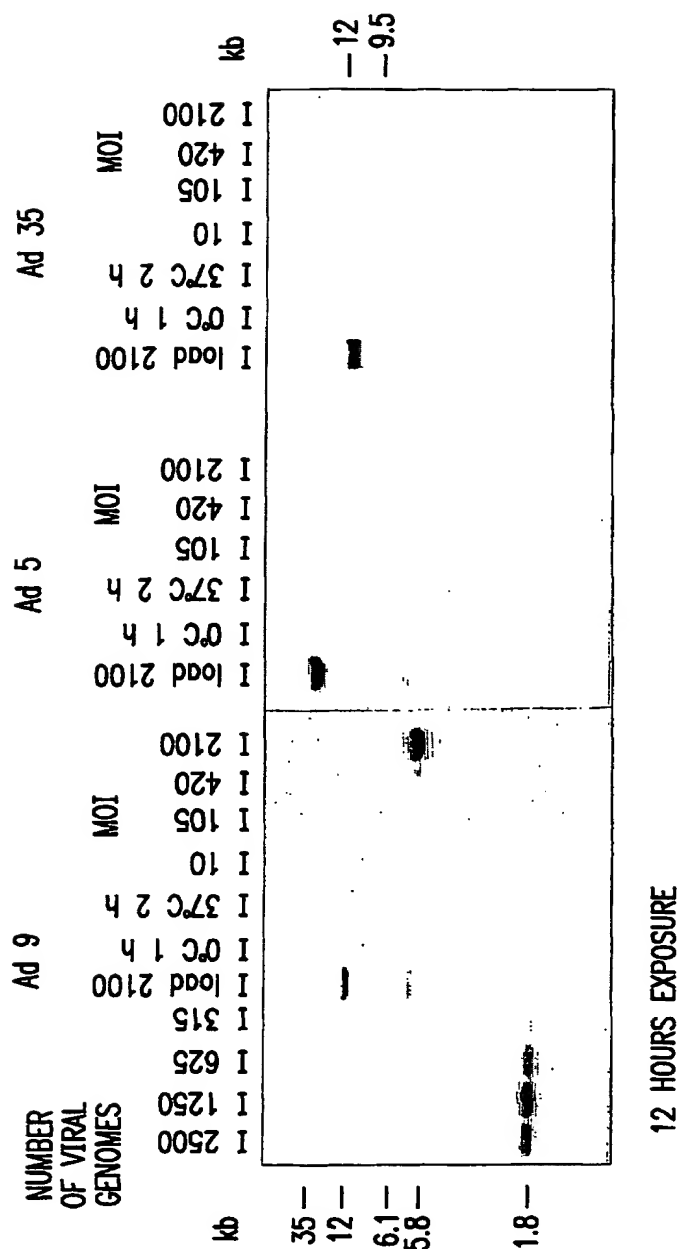


FIG. 15A



17/34

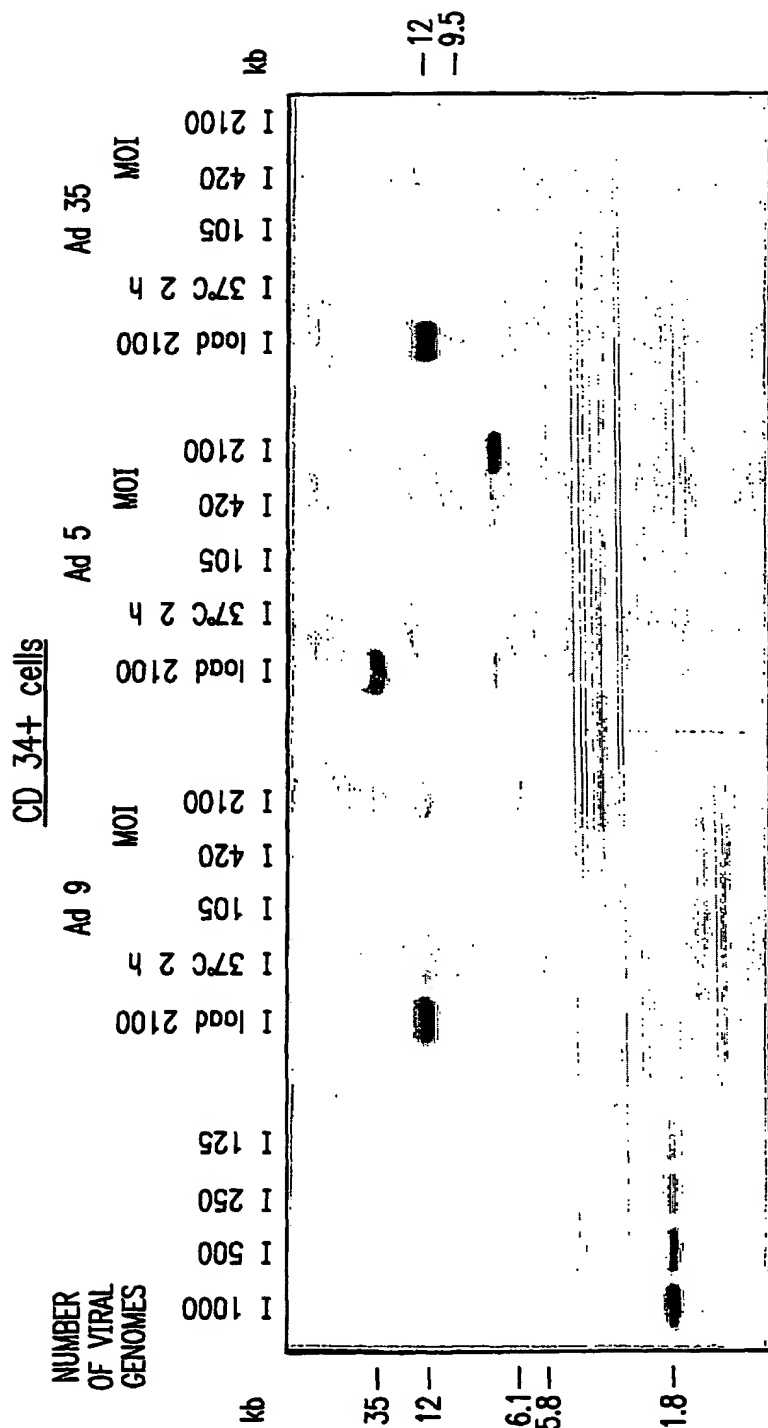


FIG. 15B

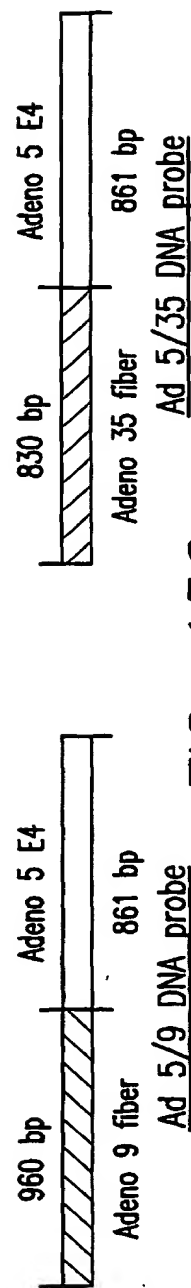


FIG. 15C

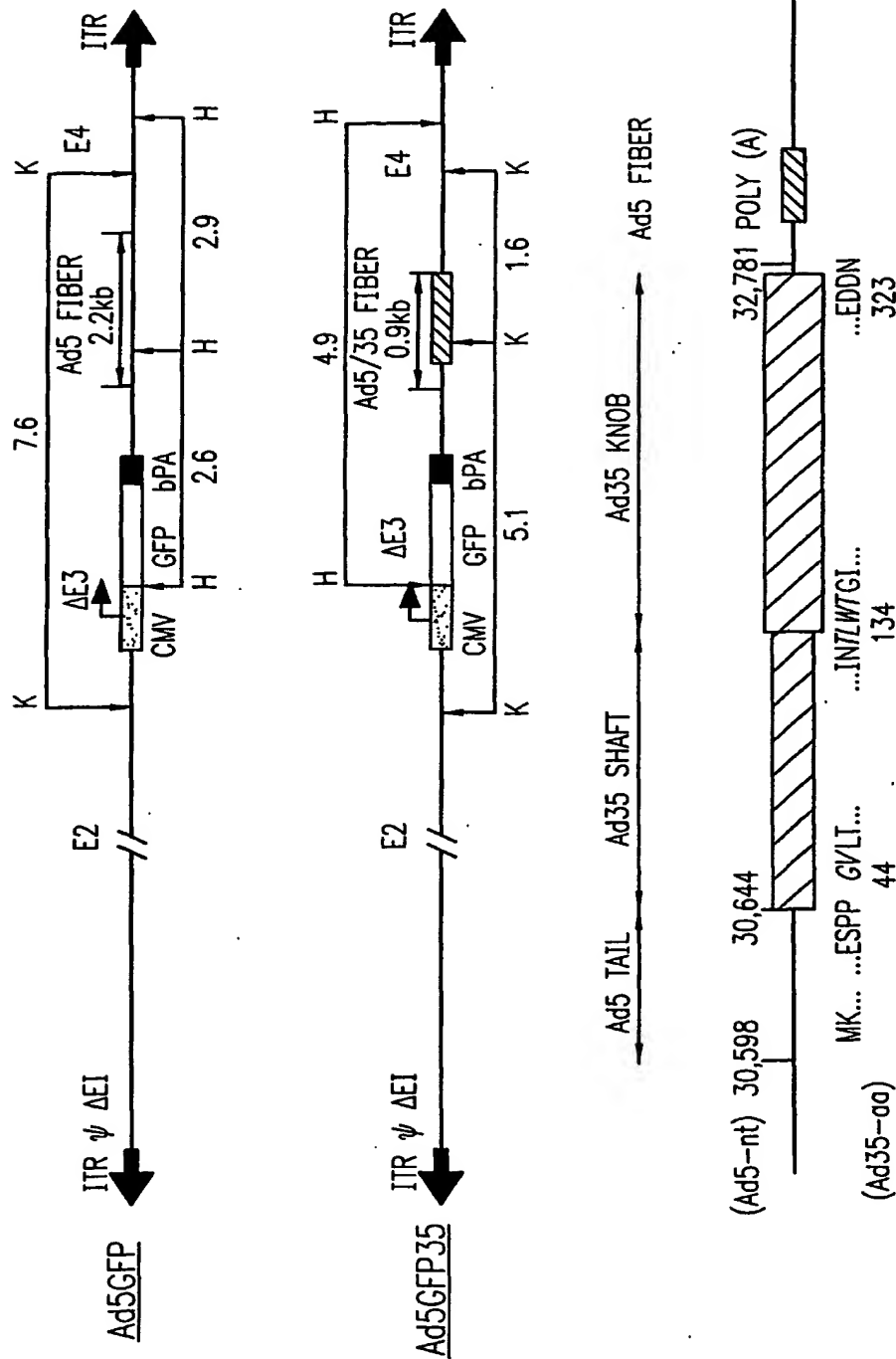


FIG. 16A

19/34

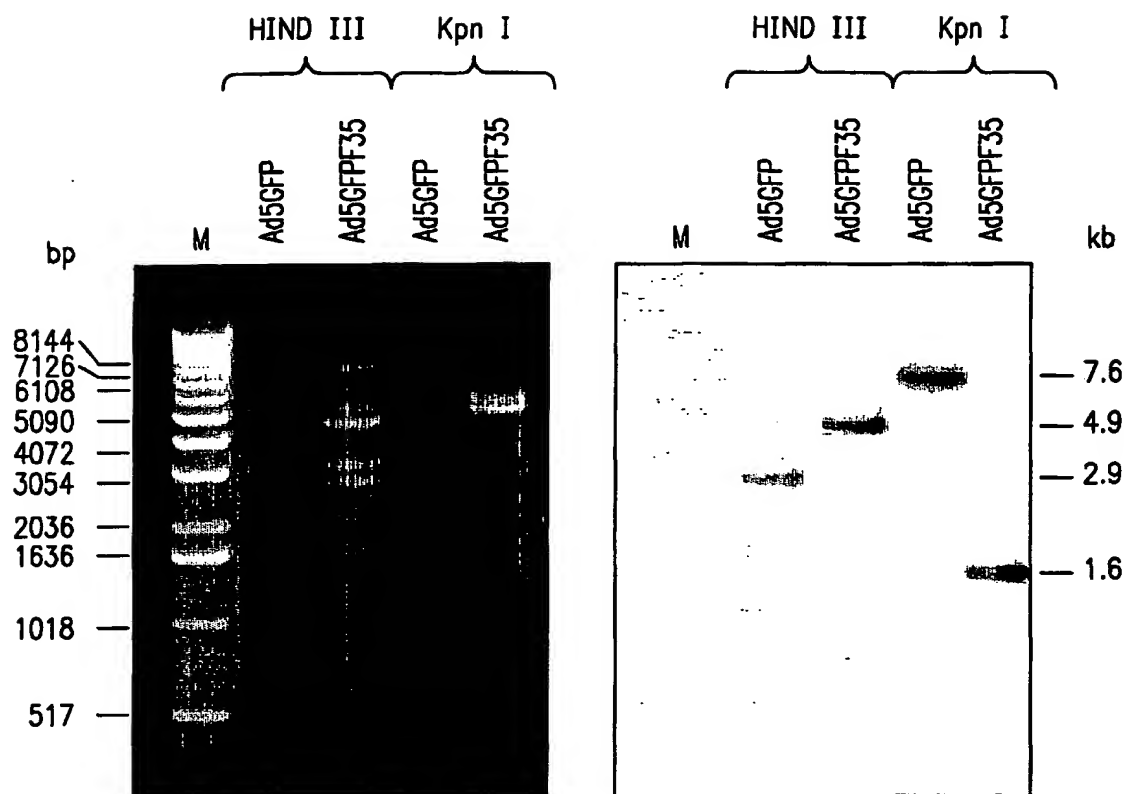


FIG. 16B

20/34

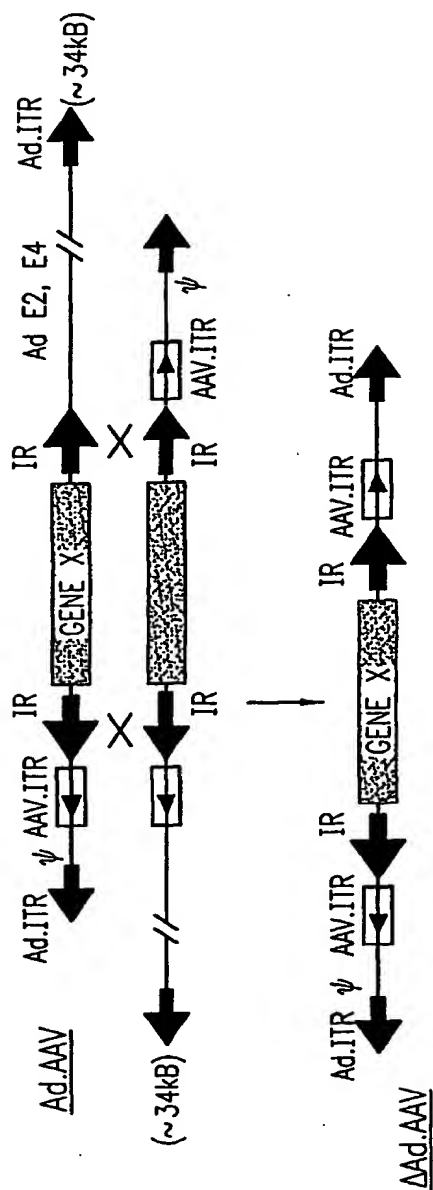


FIG. 17A

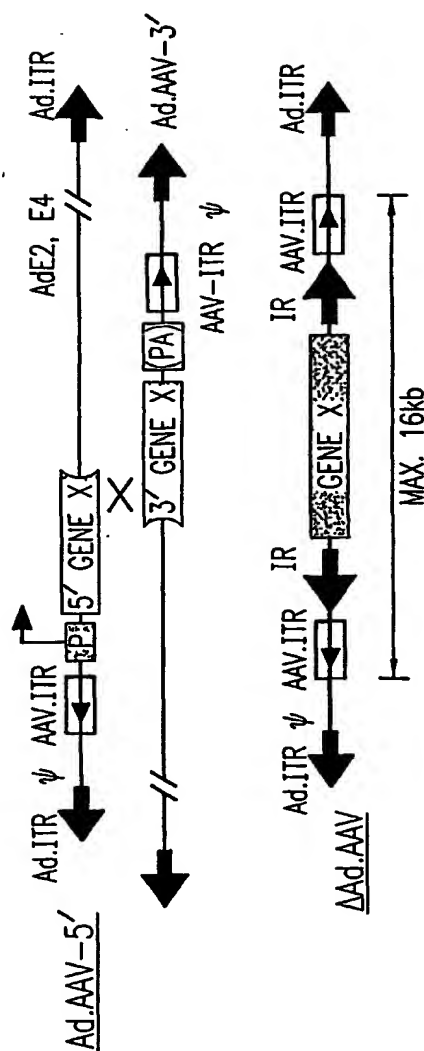


FIG. 17B

21/34

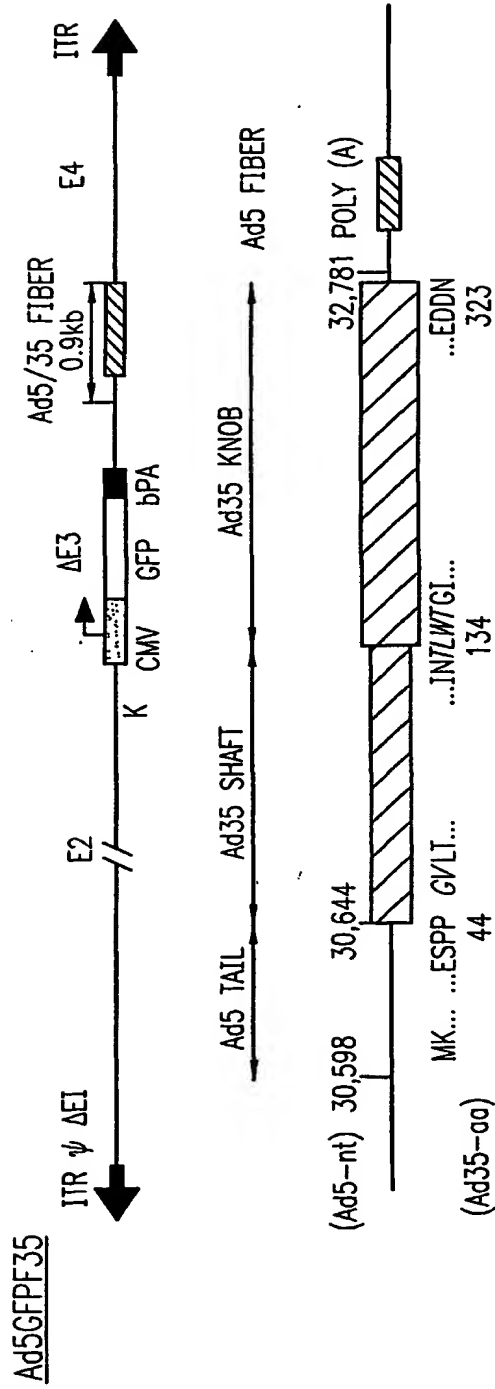


FIG. 18

22/34

■ CONTROL    □ Ad5GFP    ▨ Ad5GFPF35    ▩ Ad 35    ▪ Ad 3

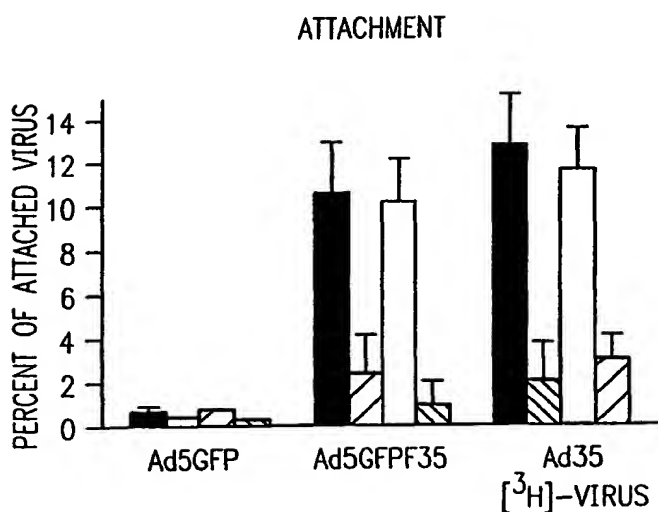


FIG. 19A

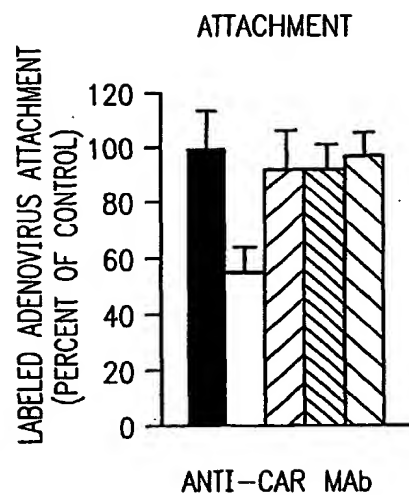


FIG. 19B

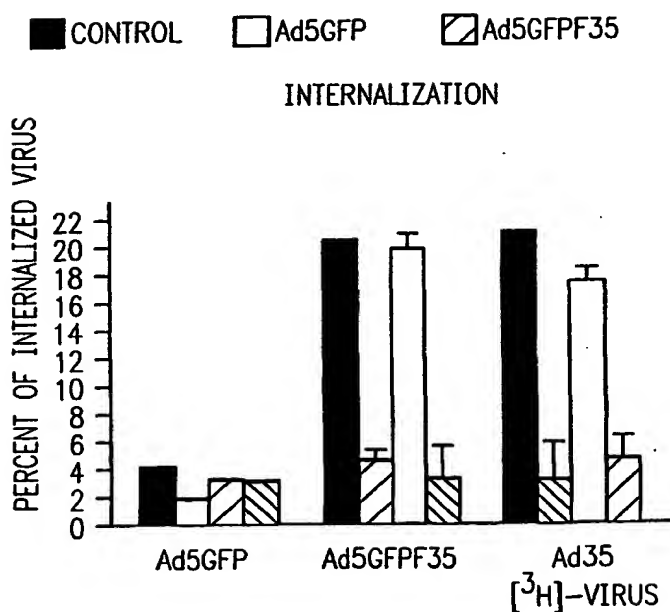


FIG. 19C

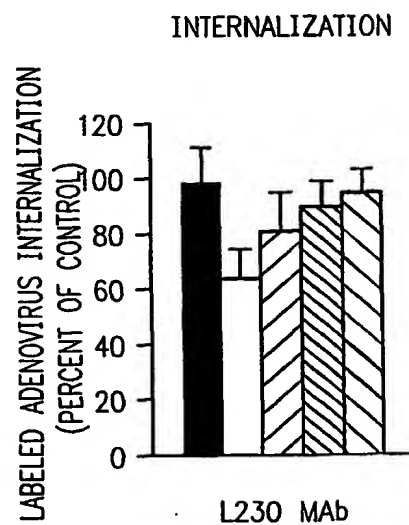


FIG. 19D

23/34

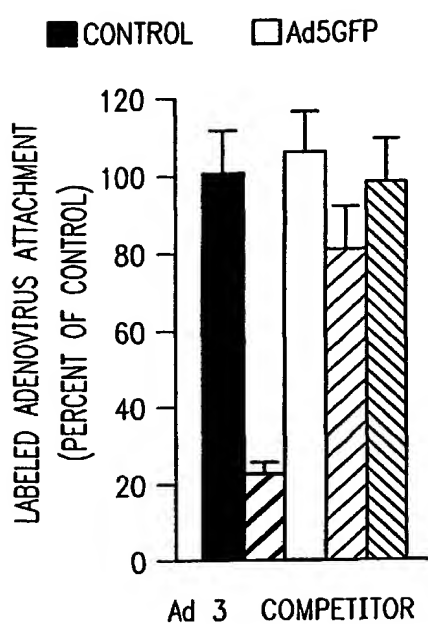


FIG. 20A

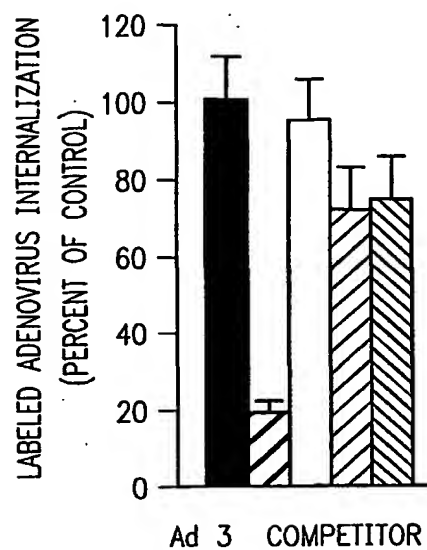


FIG. 20B

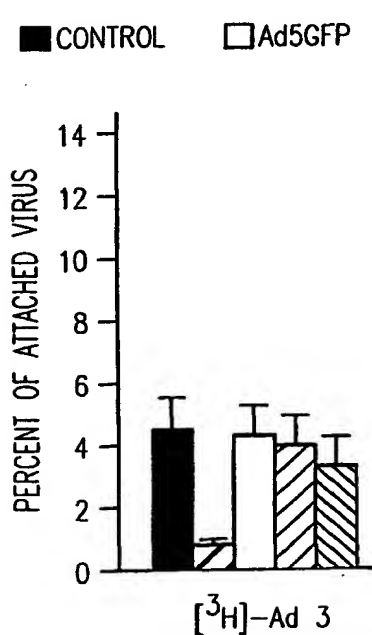


FIG. 20C

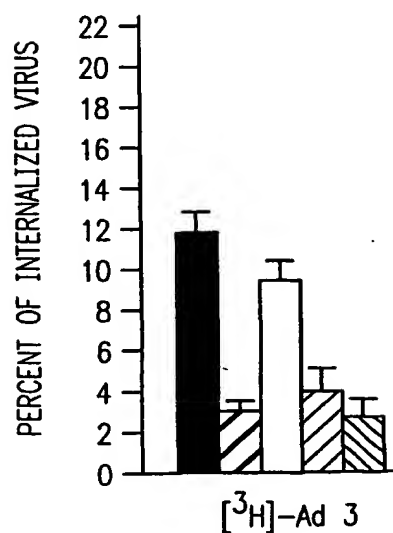


FIG. 20D

24/34

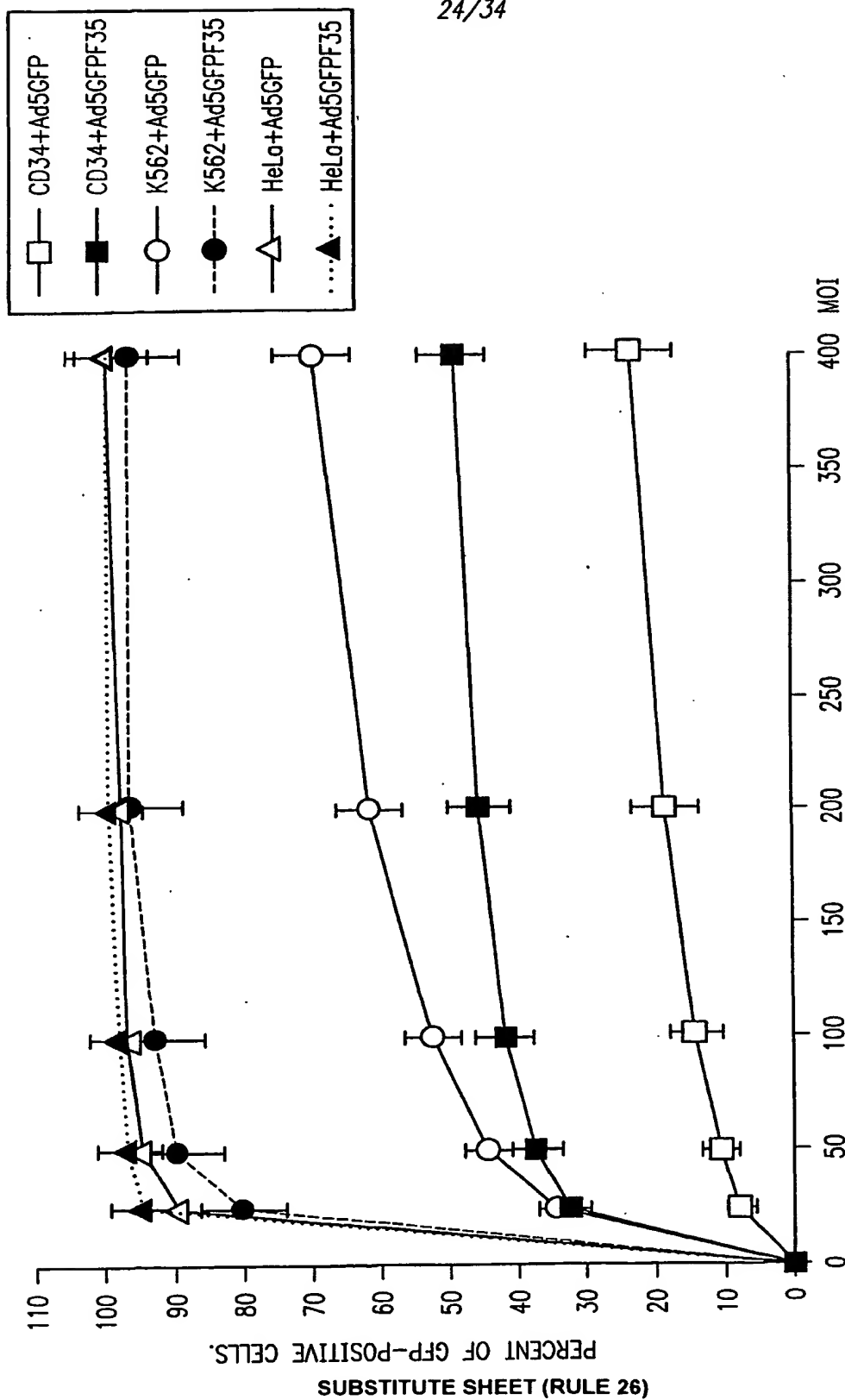


FIG. 21



25/34

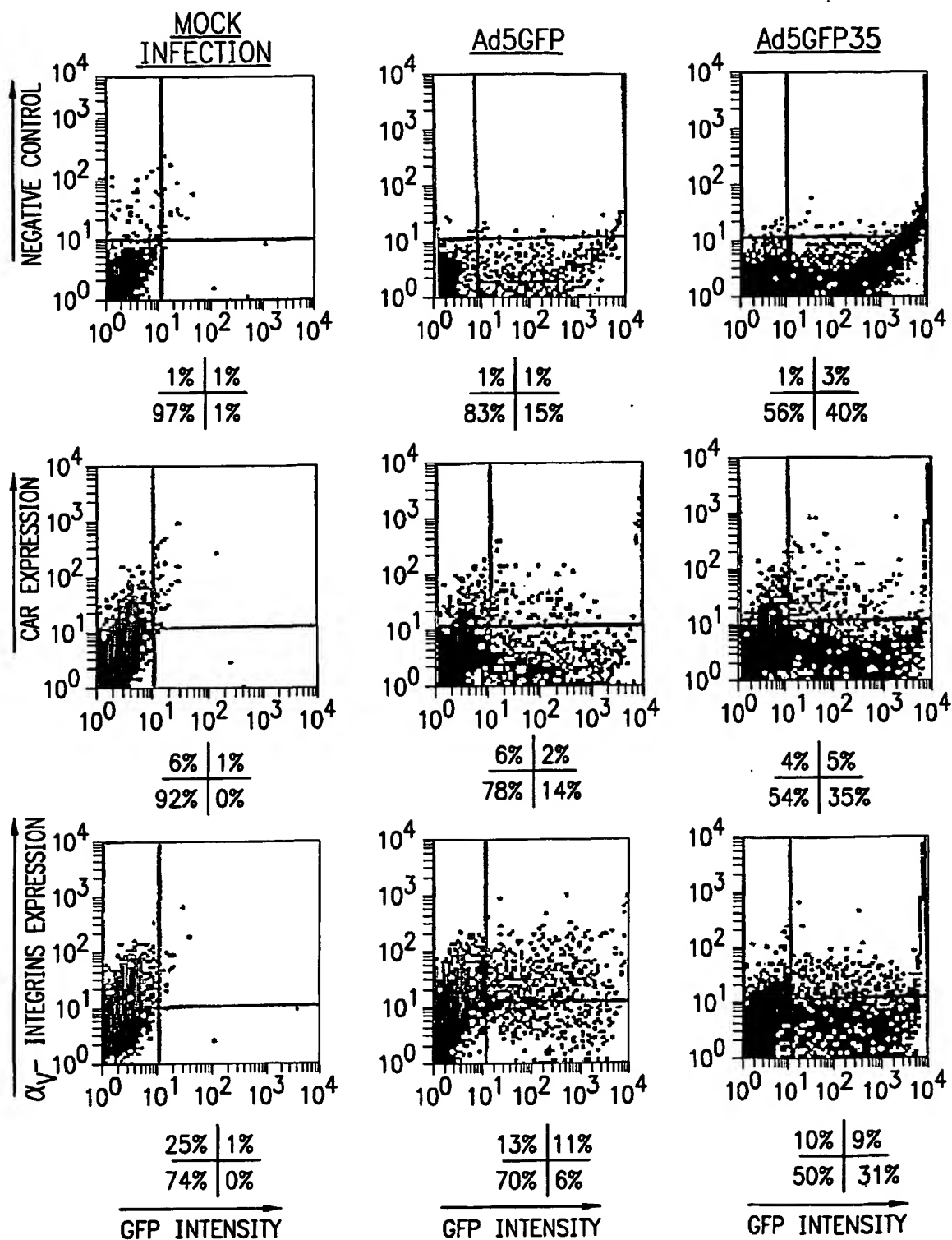


FIG. 22

SUBSTITUTE SHEET (RULE 26)

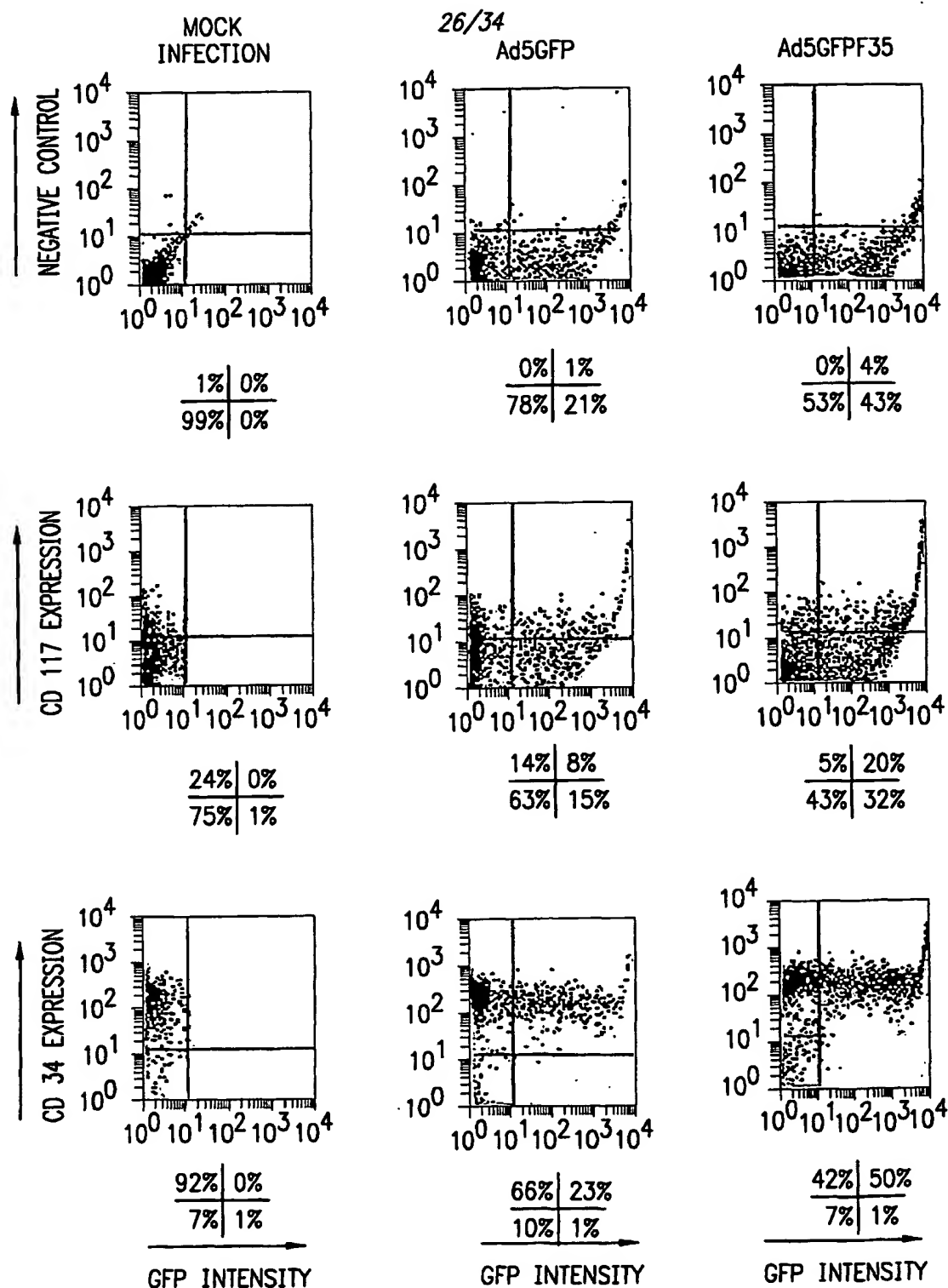


FIG. 23A

27/34

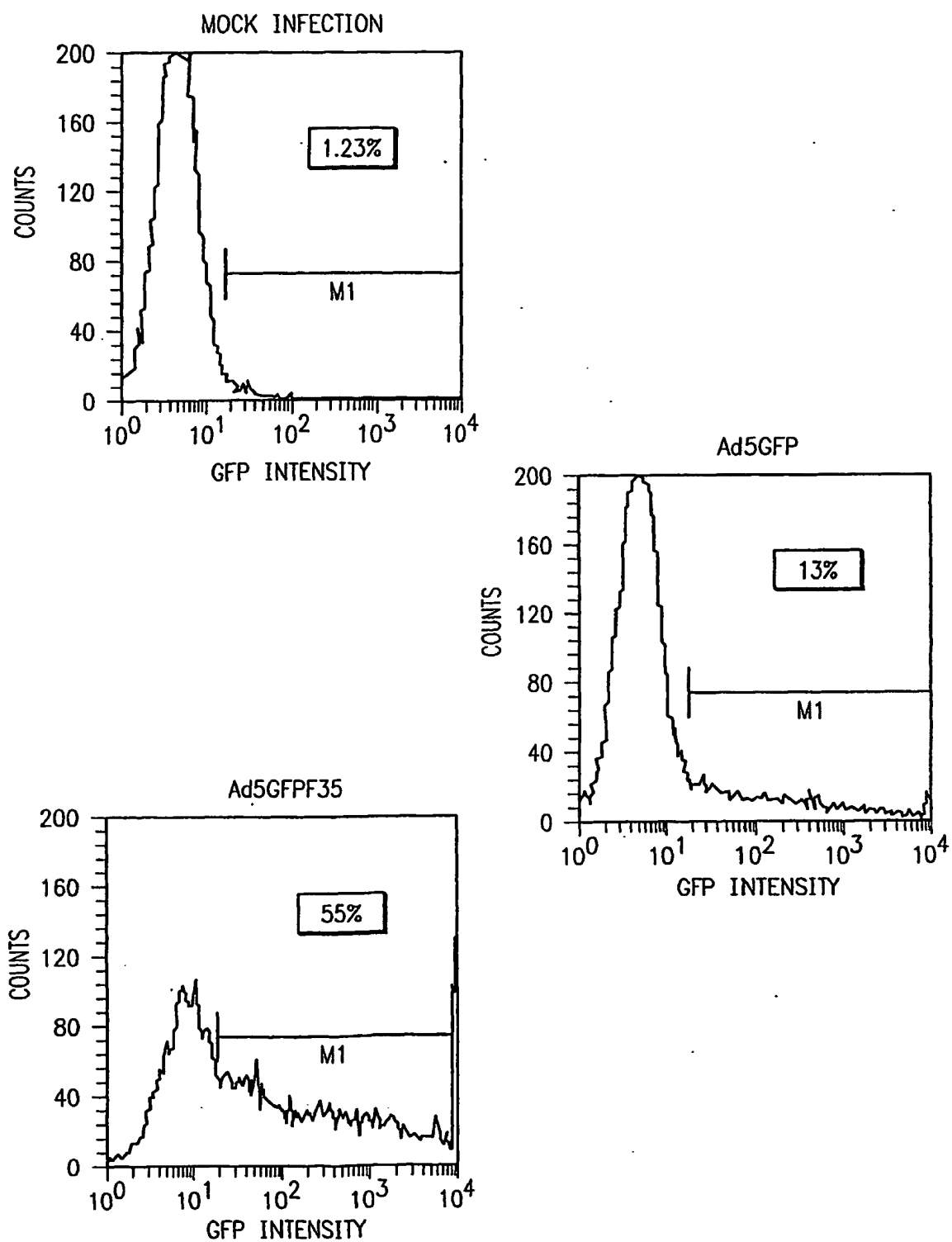
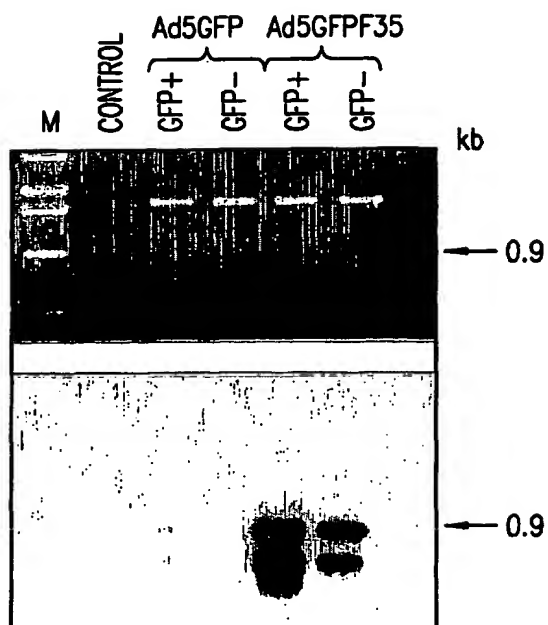
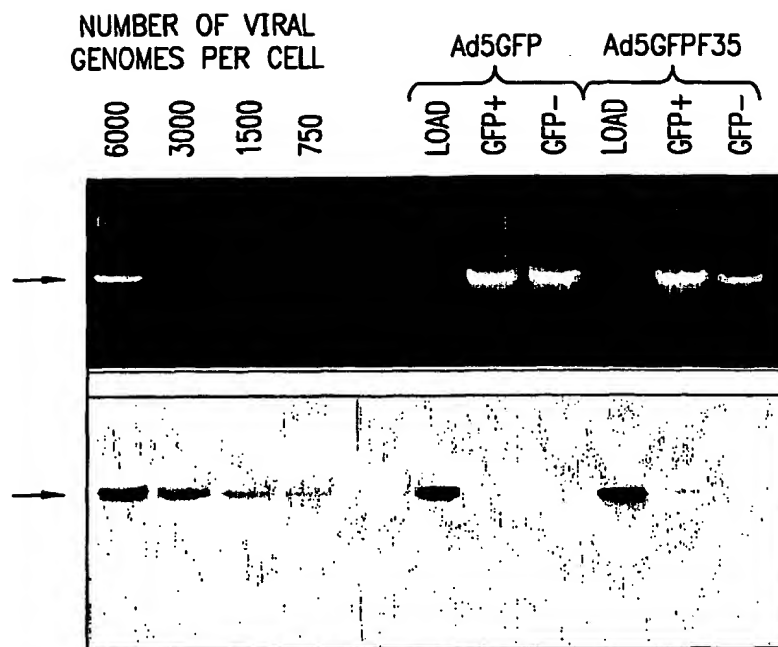


FIG. 23B

SUBSTITUTE SHEET (RULE 26)

28/34



29/34

STRUCTURE AND FEATURES OF CHIMERIC ADENOVIRUS VECTORS.

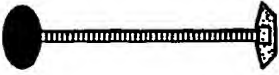


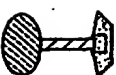
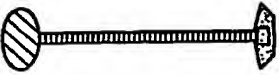

| <u>Adeno vector</u>                       |  |  |  |  |  |  |
|---|---|---|---|--|---|---|
|   | Ad 5L   | Ad 5S   | Ad 5/9L   | Ad 5/9S  | Ad 5/35L  | Ad 5/35S  |
| <u>Binding to CAR</u>                     | YES   | YES   | YES   | YES  | NO  | NO  |
| <u>genome/PFU ratio</u>                   | 13  | 143   | 29  | 582  | 17  | 14  |
| <u>Attachment to/internalization into</u> |   |   |   |  |   |   |
| 293                                       | +++   | ++  | ++  | +/-  | +++   | +++   |
| K562                                      | -/+   | --/+  | --/+  | ---  | +++   | +++   |
| Y79                                       | ++  | +   | +   | +/-  | +   | +   |
| <u>Gene transfer efficiency</u>           |   |   |   |  |   |   |
| 293                                       | +++   | ++  | ++  | +/-  | +++   | +++   |
| K562                                      | -/+   | --/+  | --/+  | ---  | +++   | +++   |
| Y79                                       | ++  | +   | +   | +/-  | +   | +   |

FIG. 25

30/34

Top View of Adenovirus Type 5 Fiber Knob Domain

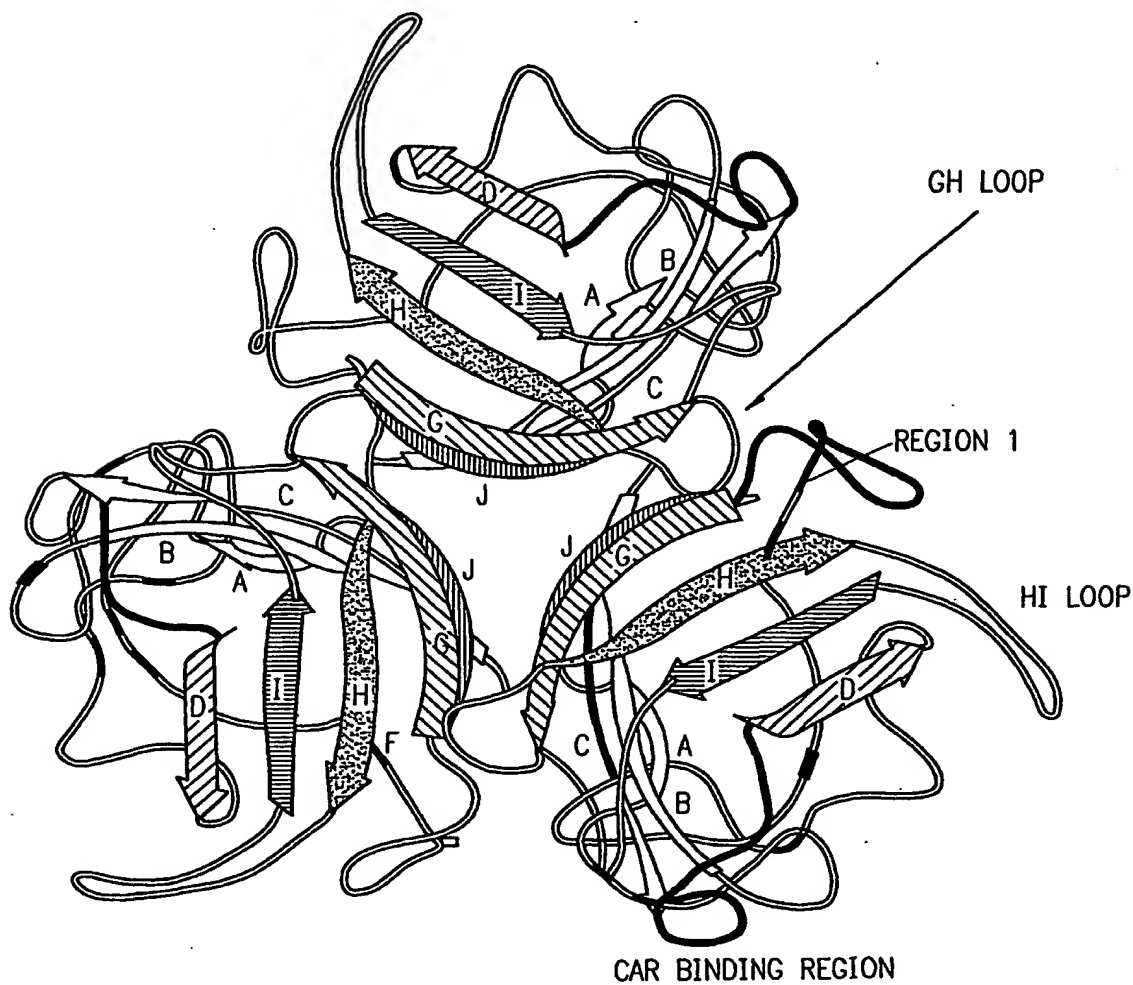
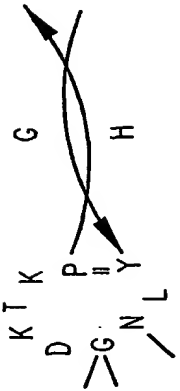


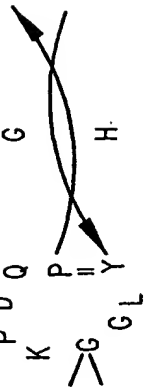
FIG. 26

Amino Acid Composition of Wt and Modified Adeno 5 GH-loop

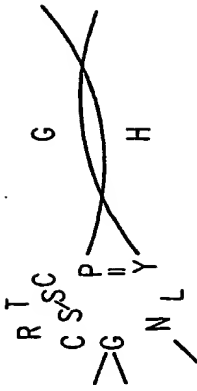
Adeno 5 wt GH-loop      V Y] L N G D K T K [P V



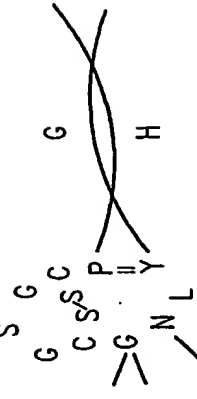
Adeno 5/9 GH-loop chimera      V Y] L G G K P D Q [P V



Adeno 5 GH-Cys 1      V Y] L N G C G S C [P V



Adeno 5 GH-Cys 2      V Y] L N G C G S G C [P V



Adeno 5 GH-peptide      V Y] L N G C G (X)<sup>10</sup> G C [P V

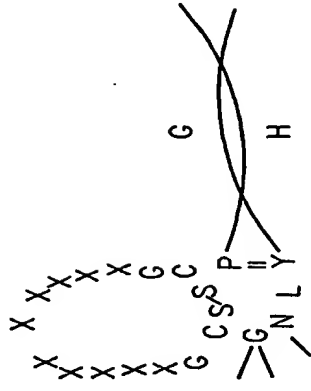


FIG. 27

32/34

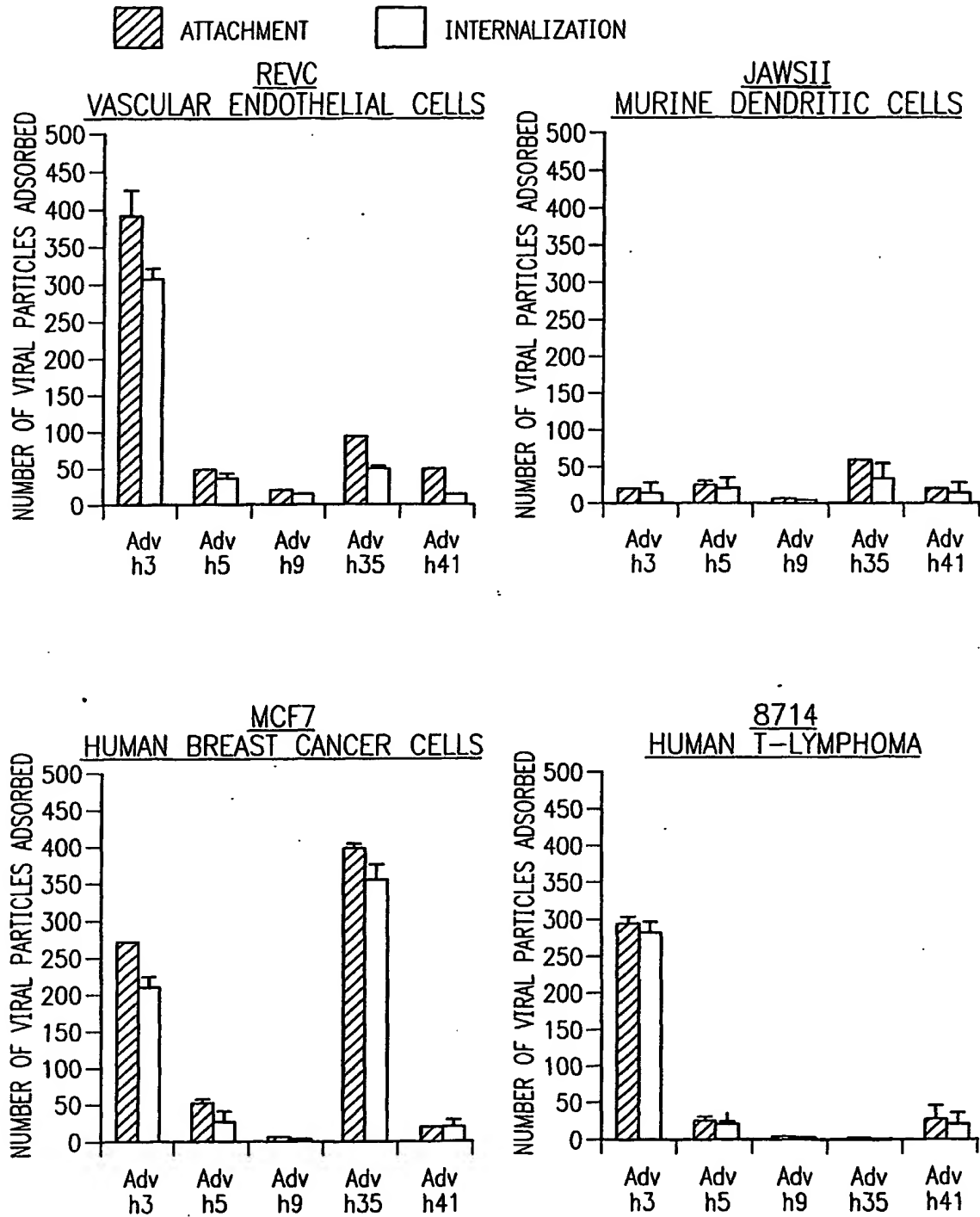


FIG. 28



33/34

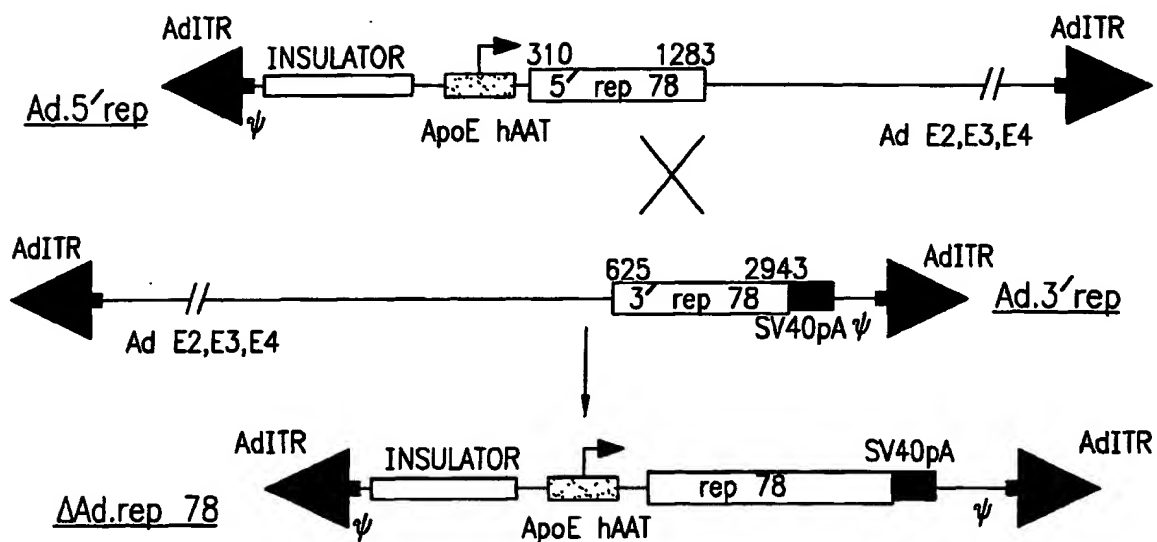


FIG. 29A

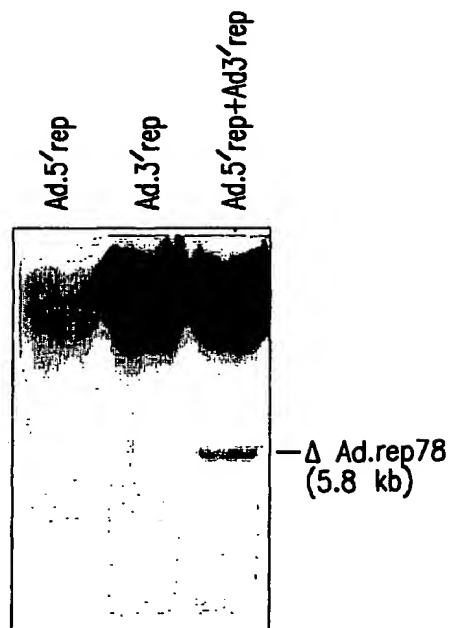


FIG. 29B

34/34

|   |   |   |   |   |   |                  |
|---|---|---|---|---|---|------------------|
| - | - | + | - | + | + | Ad-Rep5' (virus) |
| - | - | - | + | + | + | Ad-Rep3' (virus) |
| - | + | - | - | - | - | pRep78 (plasmid) |
| + | + | - | - | - | - | Ad-helper        |

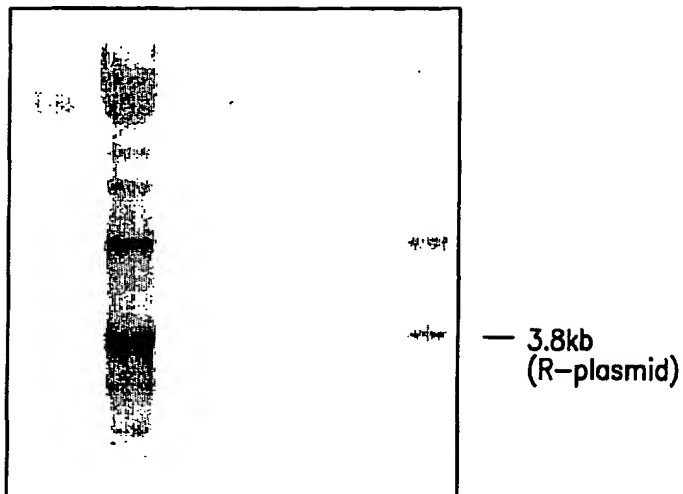


FIG. 29C

|   |   |   |   |   |                  |
|---|---|---|---|---|------------------|
| - | - | + | - | + | Ad-Rep5' (virus) |
| - | - | - | + | + | Ad-Rep3' (virus) |
| - | + | - | - | - | pRep78 (plasmid) |

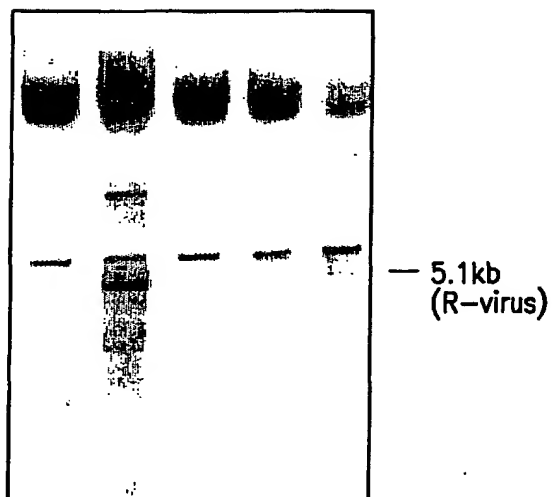


FIG. 29D

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/15442

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12N15/861 C12N15/864 C12N15/10 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                  |
|------------|---|--|
| X          | US 5 856 152 A (WILSON JAMES M ET AL)<br>5 January 1999 (1999-01-05)  | 1,2,<br>5-11,15,<br>36-44,<br>48-54,58 |
| Y          | the whole document  | 3,4,13,<br>16-32,<br>35,36,<br>46,47   |
| Y          | WO 98 54346 A (ROELVINK PETRUS W :EINFELD<br>DAVID (US); GENVEC INC (US); KOVESDI IM)<br>3 December 1998 (1998-12-03)<br>page 7, line 27 -page 14, line 26<br>the whole document<br>-/- | 3,4,<br>16-32,<br>35,56                |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*A\* document member of the same patent family

Date of the actual completion of the international search

16 November 2000

Date of mailing of the international search report

30/11/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Mandl, B

## INTERNATIONAL SEARCH REPORT

Internat'l Application No

PCT/US 00/15442

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                  |
|------------|---|--|
| X          | BURCIN M. M. ET AL.: "Adenovirus-mediated regulable target gene expression in vivo" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 96, January 1999 (1999-01), pages 355-360, XP002130326<br>ISSN: 0027-8424  | 1,2                                    |
| Y          | the whole document  | 13,46                                  |
| X          | RECCHIA A. ET AL.: "Site-specific integration mediated by a hybrid adenovirus/adeno-associated virus vector." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 96, no. 6, 16 March 1999 (1999-03-16), pages 2615-2620, XP002152917<br>ISSN: 0027-8424 | 1,2,<br>5-11,14,<br>36-44,<br>47-49,58 |
| X          | the whole document  |  |
| X          | WO 97 38723 A (BARBER JACK ;IMMUSOL INC (US); LI QI XIANG (US); YU GANG (US); YU 23 October 1997 (1997-10-23)   | 55,57                                  |
| Y          | page 4, line 14 - line 20<br>page 5, line 21 -page 6, line 10<br>page 29, line 4 -page 30, line 2<br>page 35, line 12 - line 18   | 56                                     |
| X          | WO 94 06920 A (MEDICAL RES COUNCIL ;RUSSELL STEPHEN JAMES (GB); HAWKINS ROBERT ED) 31 March 1994 (1994-03-31)<br>page 19, last paragraph  | 55,57                                  |
| A          | GALL J. ET AL.: "ADENOVIRUS TYPE 5 AND 7 CAPSID CHIMERA: FIBER REPLACEMENT ALTERS RECEPTOR TROPISM WITHOUT AFFECTING PRIMARY IMMUNE NEUTRALIZATION EPITOPES" JOURNAL OF VIROLOGY, vol. 70, no. 4, 1 April 1996 (1996-04-01), pages 2116-2123, XP000653519<br>ISSN: 0022-538X    | 16-25                                  |
| A          | the whole document  |  |
| A          | LYU Y. L. ET ALW: "Inversion/dimerization of plasmids mediated by inverted repeats." JOURNAL OF MOLECULAR BIOLOGY, vol. 285, no. 4, 29 January 1999 (1999-01-29), pages 1485-1501, XP002152918<br>ISSN: 0022-2836<br>cited in the application                                   | 1-58                                   |
|            | the whole document  |  |
|            | ---   |  |
|            | -/-   |  |

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/15442

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
|------------|--|-----------------------|
| P,Y        | STEINWAERDER D. S. ET AL.: "Generation of adenovirus vectors devoid of all viral genes by recombination between inverted repeats."<br>JOURNAL OF VIROLOGY,<br>vol. 73, no. 11, November 1999 (1999-11),<br>pages 9303-9313, XP002152919<br>ISSN: 0022-538X<br>the whole document | 1-58                  |
| P,Y        | LIEBER A. ET AL.: "Integrating adenovirus-adenovirus-associated virus hybrid vectors devoid of all viral genes."<br>JOURNAL OF VIROLOGY,<br>vol. 73, no. 11, November 1999 (1999-11),<br>pages 9314-9324, XP002152920<br>ISSN: 0022-538X<br>the whole document                   | 1-58                  |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/15442

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s)   | Publication<br>date  |
|---|---------------------|--|--|
| US 5856152 A                              | 05-01-1999          | AU 695811 B<br>AU 4405596 A<br>CA 2203808 A<br>EP 1046711 A<br>EP 0797678 A<br>JP 10507928 T<br>WO 9613598 A<br>US 5871982 A | 20-08-1998<br>23-05-1996<br>09-05-1996<br>25-10-2000<br>01-10-1997<br>04-08-1998<br>09-05-1996<br>16-02-1999 |
| WO 9854346 A                              | 03-12-1998          | AU 7604998 A<br>BR 9809173 A<br>EP 0988390 A<br>PL 337130 A  | 30-12-1998<br>01-08-2000<br>29-03-2000<br>31-07-2000   |
| WO 9738723 A                              | 23-10-1997          | AU 2678097 A<br>CA 2251738 A<br>EP 0927044 A   | 07-11-1997<br>23-10-1997<br>07-07-1999   |
| WO 9406920 A                              | 31-03-1994          | AU 689755 B<br>AU 4827893 A<br>CA 2145063 A<br>EP 0670905 A<br>JP 8504091 T<br>US 5723287 A                                  | 09-04-1998<br>12-04-1994<br>31-03-1994<br>13-09-1995<br>07-05-1996<br>03-03-1998                             |

CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 December 2000 (07.12.2000)

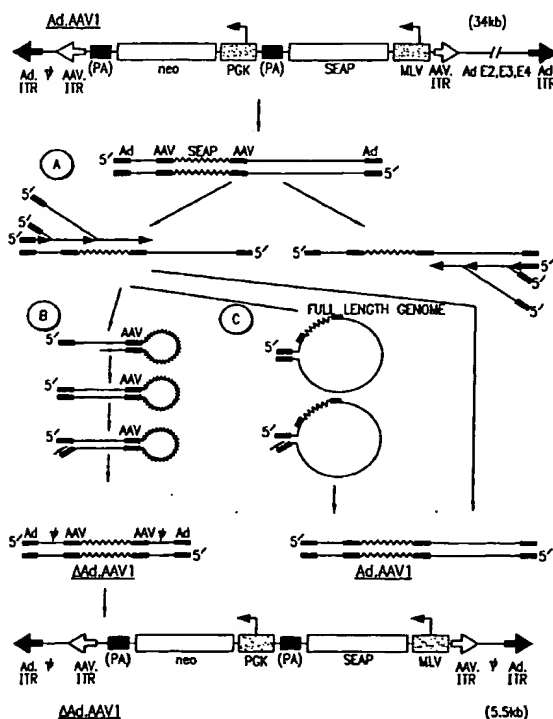
PCT

(10) International Publication Number  
WO 00/73478 A3

- (51) International Patent Classification<sup>7</sup>: C12N 15/861, 15/864, 15/10, A61K 48/00
- (21) International Application Number: PCT/US00/15442
- (22) International Filing Date: 1 June 2000 (01.06.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/137,213 1 June 1999 (01.06.1999) US  
60/161,097 22 October 1999 (22.10.1999) US
- (71) Applicant (for all designated States except US): UNIVERSITY OF WASHINGTON [US/US]; Suite 200, 1107 N.E. 45th Street, Seattle, WA 98105-4631 (US).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): LIEBER, André [DE/US]; 7251 29th Avenue, Seattle, WA 98115 (US). SHAYAKHMETOV, Dmitry [RU/US]; 6500 25th Avenue, Seattle, WA 98115 (US). FARRAR, Denise [US/US]; 1324 W. Emerson Street, Seattle, WA 98119 (US). PAPAYANNOPOULOU, Thalia [US/US]; 702 35th Avenue, Seattle, WA 98122 (US).
- (74) Agent: ADRIANO, Sarah, B.; Mandel & Adriano, Suite 60, 35 No. Arroyo Parkway, Pasadena, CA 91103 (US).
- (81) Designated States (national): AE, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID,

[Continued on next page]

(54) Title: RECOMBINANT ADENOVIRAL VECTORS EXPRESSING CHIMERIC FIBER PROTEINS FOR CELL SPECIFIC INFECTION AND GENOME INTEGRATION



(57) Abstract: The present invention provides for chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_{3/5}$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

WO 00/73478 A3



IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report

(88) **Date of publication of the international search report:**  
5 July 2001

(48) **Date of publication of this corrected version:**  
24 January 2002

(15) **Information about Correction:**  
see PCT Gazette No. 04/2002 of 24 January 2002, Section II

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



RECOMBINANT ADENOVIRAL VECTORS EXPRESSING CHIMERIC FIBER PROTEINS FOR CELL SPECIFIC  
INFECTION AND GENOME INTEGRATION5 **FIBER PROTEINS**

This invention was made, at least in part, with funding from the National Institutes of Health (Grant Nos. R01 CA 80192-01 and R21 DK 55590-01). Accordingly, the United  
10 States Government has certain rights in this invention.

**FIELD OF THE INVENTION**

This invention relates to the field of gene therapy, and in particular, to novel adenovirus  
15 (Ad) vectors that selectively infect cells for gene therapy, and to Ad vectors containing modifications of the fiber protein to allow retargeting of any adenovirus serotype.

**BACKGROUND OF THE INVENTION**

20 Gene transfer vectors require the efficient transduction of target cells, stable association with the host genome, and adequate transgene expression in the appropriate target cell, without associated toxic or immunological side effects,. Currently available viral vector systems, including recombinant retroviruses, adenoviruses and adeno-associated viruses, are not suitable for efficient gene transfer into many cell types. Retroviral vectors require  
25 cell division for stable integration. Recombinant adenoviruses are not able to infect many cell types important for gene therapy, including hematopoietic stem cells, monocytes, T- and B-lymphocytes. Moreover, recombinant adeno-associated vectors (AAV) integrate with a low frequency.

30 First generation adenoviruses have a number of properties that make them an attractive vehicle for gene transfer (Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205). These include the ability to produce purified virus at high titers in concert with highly efficient gene transfer of up to 8 kb long expression cassettes into a large variety of cell

- types *in vivo*, including non-dividing cells. Limitations of first generation adenoviruses include the development of immune responses to expressed viral proteins resulting in toxicity and virus clearance. The episomal status of adenoviral DNA within transduced cells is another limitation of first generation Ad vectors. Stable integration of adenovirus DNA into the host genome is reported only for wild-type forms of specific subtypes and appears not to occur in a detectable manner with E1/E3-deleted Ad 5 (adenovirus serotype 5) vectors widely used for gene transfer *in vitro* and *in vivo* [Hitt, M.M. et al. 1997 *Advances in Pharmacology* 40:137-205].
- 10 Recombinant AAV vectors (rAAV) integrate with a low frequency (about 1 out of 20,000 genomes) randomly as cocatemers into the host genome (Rutledge, E.A.; Russel, D.W. 1997 *J. Virology*, 71, 8429-8436). The presence of two AAV inverted terminal repeats (ITRs) and as yet unknown host cellular factors seem to be the only requirement for vector integration (Xiao, X., et al, 1997, *J. Virology*, 71, 941-948; Balague, C., et al. 15 1997, *J. Virology*, 71, 3299-3306; Yang, C.C. 1997, *J. Virology*, 71, 9231-9247). In the presence of the large AAV Rep proteins, AAV integrates preferentially into a specific site at human chromosome 19, called AAVS1 (Berns, K.I., 1996, *Fields Virology*, Fields, B.N. et al. (ed) Vol. 2, Lippincott-Raven, Philadelphia, PA, 2173-2220). The AAV capsid is formed by three coat proteins (VP1-3), which interact with specific heparin 20 sulfates on the cell surface and probably with specific receptor(s). However, many cell types, including hematopoietic stem cells, lack these structures so that rAAV vectors based on AAV2 cannot infect or transduce these cells (Malik P. et al., 1997, *J. Virology*, 71, 1776-1783; Quing, K.Y., et al. 1998, *J. Virology*, 72, 1593-1599). Other disadvantages of rAAV vectors include the limited insert size (4.5-5kb) that can be 25 accommodated in rAAV vectors lacking all viral genes and low transducing titers of rAAV preparations.
- Adenovirus infection is initiated by attaching to the cell surface of Ad 5 via its fiber protein (for a review, see Shenk, T. 1996 *Fields Virology*, Vol. 2, Fields, B.N. et al. (ed) 30 Vol. 2, Lippincott-Raven, Philadelphia, PA, 2111-2148). The distal, C-terminal domain of the trimeric fiber molecule terminates in a knob, which binds to a specific cellular

receptor identified recently as the coxackie-adenovirus receptor (CAR) (Bergelson, J.M. et al. *Science*, 275, 1320-1323). After binding, in an event independent of virus attachment, Arg-Gly-Asp (RGD) motifs in the penton base interact with cellular integrins of the  $\alpha 3$  and  $\beta 5$  types. This interaction triggers cellular internalization whereby the virion achieves localization within the endosome. The endosomal membrane is lysed in a process mediated by the penton base, releasing the contents of the endosome to the cytoplasm. During these processes, the virion is gradually uncoated and the adenoviral DNA is transported to the nucleus where replication takes place. The terminal protein, which is covalently attached to the viral genome and the core protein V that is localized on the surface of the cores have nuclear localization signals (NLSs) (van der Vliet, B. 1995, *The Molecular Repertoire of Adenoviruses*, Vol. 2, Doerfler, W. and Boehm, P.(ed.), Springer Verlag, Berlin, 1-31). These NLSs play a crucial role in directing the adenoviral genome to the nucleus and probably represent the structural elements which allow adenovirus to transduce non-dividing cells. When the double-stranded, linear DNA reaches the nucleus, it binds to the nuclear matrix through its terminal protein.

Since the cell types that can be infected with Ad5 or Ad2 vectors are restricted by the presence of CAR and specific integrins, attempts have been made to widen the tropism of Ad vectors. Genetic modification of adenovirus coat proteins to target novel cell surface receptors have been reported for the fiber (Krasnykh, V. et al. 1998 *J. Virology*, 72, 1844-1852, Krasnykh, V. et al. 1996 *J. Virology*, 70, 6839-6846, Stevenson, S.D., et al. 1997, *J. Virology*, 71, 4782-4790), penton base (Wickham, T.J., et al. 1996, *J. Virology*, 70, 6831-6838; Wickham, T.J., et al. 1995, *Gene Therapy*, 69, 750-756), and hexon proteins (Crompton, J., et al. 1994, *J. Gen. Virol.* 75, 133-139). The most promising modification seems to be the functional modification of the fiber protein or more specifically of the fiber knob as the moiety, which mediates the primary attachment. Two groups have reported the generation of fibers consisting of the Ad5 tail/shaft and the knob domain of Ad3 (Krasnykh, V. et al. 1996 *supra*, Stevenson, S.D., et al. 1997, *supra*). Recently, recombinant adenoviruses with fibers containing C-terminal poly-lysine, gastrin-releasing peptide, somatostatin, E-selectin-binding peptide, or oligo-Histidines were produced in order to change the native tropism of Ad5. Krasnikh et al. found (Krasnykh, V. et al.

1998 *supra*) that heterologous peptide ligands could be inserted into the H1 loop of the fiber knob domain without affecting the biological function of the fiber. Based on studies with other Ad serotypes, it appears that the length of the fiber shaft is a critical element, determining the efficiency of interaction with cell surface integrins and the internalization process. Thus far, there is no reported data demonstrating successful retargeting of Ad5 vectors for a specific cell type.

Therefore, there is a present need for an improved adenovirus vector which can be targeted efficiently to a variety of cell types and tissues and remain stably integrated in the host genome with minimal antigenicity to the host. The present invention discloses novel chimeric adenoviral (Ad) Ad-AAV vectors, which express a modified fiber protein on their capsid, for specifically targeting the vector. Methods of making, uses and advantages of these vectors are described. In addition, the alteration described for the knob and shaft domains of the fiber protein provide a novel approach to retarget any adenovirus serotype for cell specific infection.

#### SUMMARY OF THE INVENTION

The present invention provides for novel chimeric Ad- vectors carrying transgene, or portions of transgenes for stable and efficient gene transfer into diverse cell types or tissues in a CAR- and/or  $\alpha_v\beta_{3/5}$ - independent manner. Also provided are methods for producing such vectors and the use thereof for gene therapy to target a specific cell type or tissue.

The recombinant adenovirus vectors of the invention (Example I) provide a novel design that allows for the easy production and delivery of a "gutless" adenoviral vector with the added advantage of stable integration of the transgene into the host genome of different cell type. The adenoviral vector described is devoid of all adenoviral sequences except for the 5' and 3' cis elements necessary for replication and virion encapsidation. The adenovirus-associated virus sequences of the invention comprising the 5' (right) and 3' (left) inverted terminal repeats (ITRs) flank the transgene gene cassette such that they

direct homologous recombination during viral replication and viral integration into the host genome. In one embodiment AAV-ITR flanking sequences are used. The vector also contains a selected transgene(s) operably linked to a selected regulatory element and a polydenylation stop signal, which is in turn flanked by the flanking sequences described above. The selected transgene(s) can be linked under the same regulatory elements or under separate regulatory elements in the same orientation or in opposite orientations with respect to each other. The selected transgene(s) are any gene or genes which are expressed in a host cell or tissue for therapeutic, reporter or selection purposes. This vector is characterized by high titer transgene delivery to a host cell and the ability to stably integrate the transgene into the host genome. Also provided is a method to improve the integration frequency and site specific integration by incorporating an AAV rep protein into the recombinant hybrid vector.

The invention also provides chimeric fiber proteins (Example II), which includes naturally occurring fiber proteins in which a portion or portions of the sequence are modified to alter cell or tissue specificity of infection. Altered fiber protein sequences can include fiber protein domains (the knob domain, the shaft domain, and the tail domain) from other or the same adenovirus serotypes or from randomly selected peptides. A chimeric fiber protein can be entirely composed of non-naturally occurring sequences. The invention further relates to nucleic acid sequences encoding the chimeric fiber proteins. These nucleic acid sequences can be naturally occurring, a mixture of naturally occurring and non-naturally occurring sequences, or entirely non-naturally occurring sequences.

The heterologous fiber protein sequences described herein can be inserted into any adenovirus based vector which contains a capsid, rendering the virus capable of specifically infecting a given cell or tissue. Adenoviral vectors having such a heterologous fiber sequence can be used to direct gene transfer into desired cells. For stable integration of the transgene cassette into the host genome, the chimeric Ad-AAV vector described in the invention is the preferred vector of use.

The invention also includes a library of adenoviruses displaying random peptides in their fiber knobs can be used as ligands to screen for an adenovirus variant with tropism to a particular cell type *in vitro* and *in vivo*.

- 5 The chimeric Ad- vectors described herein include the Ad.AAV genome with a modified fiber protein expressed on its capsid. These chimeric vectors are designed to infect a wide variety of cells, in particular, the cells which can only be poorly transduced by the commonly used retroviral, AAV and adenoviral vectors. These cells include, but not limited to, hematopoietic stem cells, lung epithelial cells, dendritic cells, lymphoblastoid  
10 cells, and endothelial cells. Hematopoietic stem cells such as CD34+ cells can be targeted for gene therapy of sickle cell anemia and thalassemia using the vector described herein. The chimeric Ad-AAV vector capable of transducing genes into endothelial cells can be used in gene therapy for vascular diseases such as atherosclerosis or restinosis after coronary artery surgery.

15

#### **BRIEF DESCRIPTION OF THE FIGURES**

Figs. 1A-1C display a proposed mechanism for forming of  $\Delta$ Ad.AAV1 genome.

- 20 Figs. 2A and 2B show electron photomicrographs of hybrid virus particles: Fig. 2A shows Ad.AAV1 and Fig. 2B shows  $\Delta$ Ad.AAV1.

- Fig. 3 illustrates analysis of  $\Delta$ Ad.AAV1 genomes after transduction of SKHep1 cells. Pulse field gel eletrophoresis (PFGE).  $1 \times 10^6$  control Sk Hep1 cells (SKHep1) (lanes 1-3, 5, 9). SKHep1 cells from G418 resistant pools ( $\Delta$ Ad.AAV1) (infected with  $\Delta$ Ad.AAV1 and selected for 4 weeks) (lanes 6-8, 10-12), or SKHep1 cells collected at 3 days after  
25 infection with 2000 genomes Ad.AAV1 (Ad) lanes 4, 13) are sealed in agarose plaques, lysed *in situ* and subjected to PFGE with or without prior digestion with restriction endonucleases. Southern Blot is performed with a SEAP specific probe. U = undigested,  
30 P = digested with PI- Scl1, I = I-CeuI, E = EcoRI.

Figs. 4A and 4B show response of K562 and CD34+ cells respectively after infection with  $\Delta$ Ad.AAVBG. Cells are incubated for 6 hours with virus under agitation. At day 3 after infection, transduction frequency is calculated based on the number of X-Gal positive cells. Viability is tested by trypan blue exclusion. N = 3, SEM < 10%.

5

Fig. 5 shows Rep expression in SKHep1 and 293 cells after plasmid transfection.  $5 \times 10^5$  cells are transfected with pAAV/Ad, pRSVrep, or pPGKrep by Ca-phosphate co-precipitation. Three days after transfection, cells are harvested. Lysates are separated on a 10% PA gel, followed by Western Blot with Rep specific antibodies (03-65169),  
10 American Research Products), and developed with ECL (Amersham).

Fig. 6 shows detection of vector integration into AAVS1 by PFGE.

Fig. 7 shows strategy for creating an  $\Delta$ Ad.AAV hybrid vector capable of site-specific  
15 integration. Arrows indicate promoters, (PA) = polyadenylation signal.  $\Psi$  = adenoviral packaging signal.

Fig. 8A-8B shows vectors for transduction studies with SNori as expression unit and analysis of vector integration on genomic DNA from a small cell number. Analogous  
20 vector sets can be generated with  $\beta$ -galactosidase (BG) or green fluorescence protein (GFP) as reporter genes.

Fig. 9 shows strategy for substituting the Ad5 fiber sequence by the heterologous fiber X genes using recombination in *E. coli*.

25

Fig. 10 shows the expression of CAR and  $\alpha_v$ -integrins on test cells. For flow cytometry analysis, HeLa, CHO, K562, and CD34+ cells were incubated with monoclonal anti-CAR (RmcB, 1:400 dilution) or anti- $\alpha_v$ -integrin antibodies (L230, 1:30 dilution). As a negative control, cells were incubated with an irrelevant mouse monoclonal antibody (anti-BrdU,  
30 1:100 dilution). The binding of primary antibodies was developed with anti-mouse IgG-

FITC labeled conjugates (1:100 dilution). Data shown represent the average results of quadruplicate analyses performed on  $10^4$  cells.

Fig. 11 shows the electron microscopy of adenovirus particles. Purified particles from Ad5, 9, and 35 were negative contrast stained and analyzed at a magnification of 85,000x. Defective particles are highlighted by arrows.

Fig. 12 shows the analysis of attachment and internalization of different serotypes to CHO, HeLa, K562, and CD34+ cells. Equal amounts of [ $^3$ H]-thymidine-labeled virions of Ads 3, 4, 5, 9, 35, and 41 (measured by OD<sub>260</sub>, and equivalent to an MOI of 400 pfu per cell for Ad5) were incubated for one hour on ice as described in Materials and Methods. Cells were then washed, and the number of labeled virions bound per cell was determined. For internalization studies, viruses were first allowed to attach to cells for 1 h on ice. Then, unbound viral particles were washed out. Cells were then incubated at 37°C for 30 min followed by treatment with trypsin-EDTA and washing to remove uninternalized viral particles. The data were obtained from two to four independent experiments performed in triplicate. Note the different scale on the Y-axes for CD34+ cells.

Figs. 13A-13C show attachment and internalization of different adenovirus serotypes to Hela, CHO and 293 cells respectively.

Figs. 14A and 14B show attachment and internalization of different adenovirus serotypes to CD34+ and K-562 cells respectively.

Fig. 15A-15C shows the analysis of viral replication in K562 and CD34+ cells by Southern blot analysis of methylated viral DNA. Replication studies were performed with  $1 \times 10^5$  K562 cells (A) or CD34+ cells (B), infected with methylated Ad5, Ad9 or Ad35. The lane labeled as "load" represents DNA that was extracted from the media/cell mixture immediately after adding the indicated viral dose to cells. The intensities of bands corresponding to methylated and un-methylated viral DNA indicate that ~85% of



the input virus was methylated. To quantify adsorption and internalization, DNA analysis was performed after prior incubation of virus with cells at 0°C (adsorption) or 37°C (internalization). For dose dependent replication studies, the indicated viral dose (expressed as the number of genomes) was added to the cells, and cellular genomic DNA together with viral DNA was extracted 16 hours or 36 hours post-infection for K562 and CD34+ cells, respectively. Identical amounts of sample DNA were analyzed by Southern blot. For quantification purposes, Ad9 replication was analyzed together with Ad5 using an Ad5/9 chimeric probe that hybridizes with DNA of both serotypes (C). The analysis of Ad5 versus Ad35 replication was performed with the corresponding Ad5/35 chimeric probe. Since separate hybridizations with both Ad5/35 and Ad5/9 probes gave identical signal intensities for Ad5 DNA only one panel is shown for Ad5 replication in test cells. To produce distinguishable fragments specific for the methylated or non-methylated status of viral genomes, Ad5 DNA was digested with *Xho* I, while Ad9 and Ad35 DNA was digested with *Xho* I and *Hind* III. The bands specific for methylated (not-replicated) viral DNA were ~12kb for Ad9, 35kb for Ad5, and ~12kb for Ad35. The fragments specific for non-methylated DNA were 5.8kb for Ad9, 6.1kb for Ad5, and 9.5kb for Ad35. Chimeric Ad5/9 and Ad5/35 DNA fragments (1.8kb) were used as quantification standards and applied onto gel together with digested viral/cellular DNA (shown on the left part of the figures).

Fig. 16A-16B shows the structure of Ad5GFP and chimeric Ad5GFP/F35 vectors. A) Schematic diagram of the original E1/E3 deleted Ad5-based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes. A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The

Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal. The region of Ad5GFP/F35 encoding for chimeric fiber was completely sequenced with Ad5 specific primers (see Material and Methods). B) Restriction analysis of viral genomes. Viral DNA was isolated from purified Ad5GFP and Ad5GFP/F35 particles as described elsewhere. One microgram of DNA was digested with *Hind* III or *Kpn* I and separated in ethidium bromide stained agarose gels (left panel) which were subsequently blotted and analyzed by Southern blot with an Ad5 E4 specific probe (nt 32,775-33,651) (right panel). Specific patterns, designating the correct structure for both viral vectors were detected. The *Hind* III fragments specific for Ad5GFP and Ad5GFP/F35 were 2.9kb and 4.9kb, respectively. The *Kpn* I fragment that confirmed the correct Ad5GFP/F35 structure was 1.6kb compared to a 7.6kb Ad5GFP fragment. M - 1kb ladder (Gibco-BRL, Grand Island, NY).

Fig. 17 shows the generation of  $\Delta$ Ad.AAV genomes by recombination between inverted homology regions. Recombination between two inverted repeats (IR) present in one Ad.AAV vector. The first-generation Ad.AAV vector (~34kb) contains two 1.2kb inverted homology elements flanking the transgene cassette. One AAV-ITR is inserted between the Ad packaging signal ( $\psi$ ) and the left IR. During Ad replication, recombination between the *Irs* mediates the formation of the  $\Delta$ Ad.AAV genomes with the transgene flanked by *Irs*, AAVITRs, Ad packaging signals, and Ad ITRs. These genomes are efficiently packaged into Ad capsids.

Fig. 18 shows the structure of Ad5/11, Ad5/35. Schematic diagram of the original E1/E3 deleted Ad50based vector with GFP-expression cassette inserted into the E3 region (Ad5GFP) and the chimeric vector Ad5GFP/F35 containing the Ad5/35 fiber gene. The 2.2kb Ad5 fiber gene was replaced by a 0.9kb chimeric fiber gene encoding for the short shaft and knob of Ad35 by a technique that involved PCR-cloning and recombination in *E.coli*. *Kpn* I (K) and *Hind* III (H) sites localized within or around the fiber genes are indicated. The lower panel shows the detailed structure of the chimeric fiber region. The Ad5 fiber tail [amino acids (aa): 1-44] were joined in frame to the Ad35 fiber shaft starting from its first two amino acids (GV), which are conserved among many serotypes.

A conserved stretch of amino acids TLWT marks the boundary between the last  $\beta$ -sheet of Ad35 shaft and the globular knob. The Ad35 fiber chain termination codon is followed by the Ad5 fiber poly-adenylation signal.

5 Fig. 19 shows the cross-competition for attachment and internalization of labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled viruses, and with anti-CAR or anti- $\alpha_v$ -integrins Mab. (A) For attachment studies,  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled competitor virus at 4°C for 1 h. Then, equal amounts of [ $^3$ H]-labeled viruses, at a dose equivalent to an MOI of 100 pfu per cell  
10 determined for Ad5GFP, were added to cells followed by incubation at 4°C for 1 h. Cells were then washed with ice-cold PBS, pelleted and the percentage of attached virus (cell-associated counts per minute) was determined. For analysis of cross-competition for internalization, cells were pre-incubated with a 100-fold excess of competitor virus at 37°C for 30 min before labeled virus was added. After an additional incubation at 37°C  
15 for 30 min, cells were treated with trypsin-EDTA for 5 min at 37°C, washed with ice-cold PBS, pelleted, and the percentage of internalized virus was determined. For controls, cells were incubated with labeled viruses without any competitors. Preliminary experiments had shown that the conditions chosen for competition studies allowed for saturation in attachment/internalization on K562 cells for all unlabeled competitors. (B)  $10^5$  K562 cells  
20 were pre-incubated for 1 hour at 4°C with anti-CAR MAb (RmCB, diluted 1:100) or with anti- $\alpha_v$ -integrin MAb (L230, diluted 1:30), followed by incubation with labeled viruses according to the protocols for attachment or for internalization as described above. For each particular serotype, the percentage of attached/internalized virus was compared to the control settings, where cells were preincubated under the same conditions with a  
25 1:100 dilution of an irrelevant antibody (anti-BrdU Mab) before addition of the labeled virus. Note that the specific competitors but not the corresponding controls significantly inhibited Ad5 internalization to a degree that is in agreement with published data (59).  
N>/=4

30 Fig. 20. Cross-competition for attachment and internalization of [ $^3$ H]-labeled Ad5GFP, Ad35, and chimeric Ad5GFP/F35 virions with unlabeled Ad3 virus (A), and of [ $^3$ H]-

labeled Ad3 virions with unlabeled viruses (B).  $10^5$  K562 cells were pre-incubated with a 100-fold excess of unlabeled viral particles according to attachment or internalization protocols described for Fig.6. Equal amounts of [ $^3$ H]-labeled Ad5GFP, Ad5GFP/F35, or Ad35 (A) or [ $^3$ H]-labeled Ad3 (B) were added to cells at a dose equivalent to an MOI of 100 pfu per cell for Ad5GFP. In control settings, cells were incubated with labeled viruses without any competitors. N=4.

Fig. 21 shows the transduction of CD34+, K562, and HeLa cells with Ad5GFP and chimeric Ad5GFP/F35 vectors.  $1 \times 10^5$  cells were infected with different MOIs (pfu/cell) of viruses in 100  $\mu$ l of media for 6 hours at 37°C. Virus containing media was then removed, and the cells were resuspended in fresh media followed by incubation for 18 h at 37°C. The percentage of GFP expressing cells was determined by flow cytometry. N=3

Fig. 22 shows the distribution of GFP-positive cells in subpopulations of human CD34+ cells expressing CAR or  $\alpha_v$ -integrins.  $1 \times 10^5$  CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu/cell as described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CAR (1:100 final dilution) or anti- $\alpha_v$ -integrin (1:30 final dilution) primary MAbs for 1 h at 37°C. Binding of primary antibodies was developed with anti-mouse IgG-PE labeled secondary MAbs (1:100 final dilution) at 4°C for 30 min. For each variant,  $10^4$  cells were analyzed by flow-cytometry. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The data shown were representative for three independent experiments.

25

Fig. 23A-23B shows the distribution of GFP-positive cells in a subpopulation of human CD34+ cells, expressing CD34 and CD117 (c-kit). (A) Co-localization of GFP expression with CD34 or CD117: CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell under the conditions described for Fig.8. Twenty-four hours after infection, cells were incubated with anti-CD34 PE-conjugated MAbs (final dilution 1:2) or with anti-CD117 PE-conjugated MAbs (final dilution 1:5)

30

for 30 min on ice, and  $10^4$  cells per variant were subjected to two-color flow cytometry analysis. For negative control staining, no antibodies were added to the cells before analysis. The mock infection variants represent cells incubated with virus dilution buffer only. The quadrant borders were set based on the background signals obtained with both the GFP- and PE-matched negative controls. The percentages of stained cells found in each quadrant are indicated. The experiment was performed two times in triplicates, and typically obtained results are shown. The SEM was less than 10% of the statistical average. (B) Transduction of CD34+/CD117+ cells with Ad5GFP and chimeric Ad5GFP/F35 virus vectors: CD34+ cells, cultured overnight before staining in media without SCF, were incubated with PE-labeled anti-CD117 MAb for 30 min on ice. The fraction of CD117-positive cells was sorted by FACS. More than 97% of sorted cells were positive for CD117.  $1 \times 10^5$  CD117+/CD34+ cells were infected with Ad5GFP or Ad5GFP/F35 at an MOI of 200 pfu per cell, as for Fig.8. Twenty-four hours post infection, the percentage of GFP positive was determined by flow cytometry. For mock infection, CD117+/CD34+ cells were incubated with virus dilution buffer only. The infections were done in triplicates, and the average percentage of GFP-expressing cells is indicated on the corresponding histogram. The SEM was less than 10% of the statistical average.

Fig. 24 shows the southern analysis of viral genomes in GFP-positive and GFP-negative fractions of CD34+ cells infected with the Ad5GFP and chimeric Ad5GFP/F35 vectors. CD34+ cells were infected with viruses at an MOI of 100 as described for Fig.21. Twenty four-hours post infection, cells were sorted by FACS for GFP positive and GFP negative fractions.  $10^5$  cells from each fraction were used to isolate genomic DNA together with viral DNA. Before cell lysis, a rigorous treatment with trypsin and DNase followed by washing was performed to exclude that genomic DNA samples were contaminated by extracellular viral DNA. A) The upper panel shows the ethidium bromide stained 1% agarose gel before blotting demonstrating that similar amounts of genomic DNA were loaded. This amount corresponded to DNA isolated from ~25,000 GFP+ or GFP- cells. The lane labeled Aload@ represents viral DNA purified from Ad5GFP or Ad5GFP/F35 virions mixed with pBluescript plasmid DNA (Stratagene) as a carrier and applied on a

gel at the amount that was actually used to infect 25,000 cells. As a concentration standard, a serial dilution of Ad5GFP genomes was loaded on the gel (left side). For Southern analysis (lower panel), an 8kb-long HindIII fragment corresponding to the E2 region of Ad5 was used as a labeled probe. Hybridized filters were subjected to  
5 PhosphorImager analysis and then exposed to Kodak-X-OMAT film for 48 h at B70°C. The cellular/viral genomic DNA is indicated by an arrow. (B) To detect Ad5GFP genomes in transduced cells, PCR amplification followed by Southern blot hybridization was performed on the same samples that were used for quantitative Southern blot hybridization in (A). DNA purified from ~2,500 cells was subjected to PCR (95°C 1min,  
10 53°C 1min, 72°C 1min, 20 cycles with primers Ad5-F1 and Ad5-R1). One fifth of the PCR reaction was subjected to agarose gel electrophoresis (upper panel). A 0.9 kb-long DNA fragment, specific to the E4 region of Ad5 was detected for transduced Ad5GFP/F35 genomes. DNA then was blotted onto Nybond-N+ membrane and Southern blot hybridization (lower panel) with an Ad5 E4 specific DNA probe was performed. In  
15 addition to the 0.9kb DNA fragment, the PCR primers generated a smaller 0.5 kb-long fragment that also hybridized with the E4 region probe.

Fig. 25 shows the role of fiber shaft length in Ad infection strategies. CAR binding (Ad5 and Ad9) variants and Ad35, which interacts with a non-CAR receptor were analyzed on  
20 CAR expressing cells (293, Y79) and K562 cells which do not express significant CAR amounts. All vectors contain a GFP expression cassette packaged into an Ad5 capsid with modified fibers.

Fig. 26 shows the tertiary structure of Ad5 knob: localization of CAR binding sites, H-I  
25 loop and G-H loop.

Fig. 27 shows the substitution of the G-H loop with heterologous peptides.

Fig. 28 shows the attachment and internalization of metabolically labeled serotypes with  
30 human cell lines.

Figs. 29A-29D shows the generation of Rep78 expressing Ad vectors by recombination between two vectors. (A) The same strategy outlined in Fig. 15 was employed for vectors with rep 78 as a transgene. The Ad5'rep vector also contained the ApoEhAAT promoter shielded by an HS-4 insulator. The region of homology between the two fragments of the rep78 gene was 658nt in length. The Rep78 ORF was deleted for the p5 promoter. The internal Rep 40/52 start codon (at position 993) was mutated to abolish production of the small Rep proteins. Furthermore, the splice site at nt 1905 was deleted eliminating production of Rep68. The individual expression of Rep 78 was demonstrated. (B) Formation of  $\Delta$ Ad.rep78 genomes. The expected 5.8kb  $\Delta$ Adrep78 genome was only observed upon coinfection of both Ad5'rep and Ad3'rep into 293 cells as demonstrated by Southern. (C) Southern blot analysis for rescue of the recombinant AAV genome from plasmid DNA by Rep78 expressed from pCMVrep78 and  $\Delta$ Ad.rep78. The expected rescue product is 3.8kb (R-plasmid). (D) Southern blot analysis for rescue of the recombinant AAV genome from Ad.AAV viral vector genomes.

15

#### **DETAILED DESCRIPTION OF THE INVENTION**

All scientific and technical terms used in this application have meanings commonly used in the art unless otherwise specified. As used in the present invention, the following words or phrases have the meanings specified.

The term vector includes, but is not limited to, plasmids, cosmids, phagemids, and artificial chromosomes. The vector sequence may be designated as the viral "base vector" sequence. The base vector sequence is dependent upon the particular type of virus and serotype that the base vector sequence was derived from. The base vector sequence may be linked to non-vector or transgene sequences (e.g., heterologous sequences).

The transgene sequences may include sequences that confer compatibility with prokaryote or eukaryote host cells, where compatibility relates to vector replication

within a host cell. Accordingly, the transgene sequence may be a replicon sequence that directs replication of the vector within the host cell, resulting in an autonomously replicating vector. Alternatively, the transgene sequence may permit vector replication that is dependent upon the host cell's replication machinery.

5

The base vector sequences may be operatively linked to a transgene sequence that encodes a gene product, such as a polypeptide, rRNA, or tRNA. For example, the transgene may encode a polypeptide such as a viral capsid protein, or a viral fiber protein. The transgene may be derived from the same or different serotype as the base vector sequence.

10

Another example of a transgene includes a reporter gene that encodes a gene product that can be used as a selectable marker, such as drug resistance or a colorimetric marker. The reporter gene may encode a gene product which can be readily detected by, for example, a visual microscopic, immunochemical, or enzymatic assay. The preferred reporter gene encodes a gene product that can be detected by a non-destructive method that does not destroy the cell that expresses the reporter gene.

15

A therapeutic gene is another example of a transgene. A therapeutic gene encodes a gene product (e.g., polypeptide or RNA) which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or the tissue or the organ or the organism containing the host cell. The therapeutic benefit may result from modifying a function of a gene in the host genome or from the additional function provided by the therapeutic protein, polypeptide or RNA.

20

25

The base vector sequence may be linked to a transgene sequence that is a regulatory element, such as a promoter, enhancer, transcription termination signals, polyadenylation sequences. The regulatory element may direct expression of the transgene sequence that encodes a gene product by direct transcription or translation.

30 The regulatory element may regulate the amount or timing of expression of the transgene



sequence. The regulatory element may direct expression of the transgene in certain host cells or tissues (e.g., host-specific or tissue-specific expression).

5 The base vector sequence may be linked to a transgene sequence that permits the vector, to integrate into another nucleotide sequence. The integration sequence may direct integration of the whole vector or portions of the vector. The integration sequence may or may not be related to the base vector sequence. For example the integration and base vector sequences may be from the same or different viral serotype. The integration sequence may be inverted repeat sequences (ITRs) from adenovirus (Ad), adenovirus-associated virus (AAV), or HIV.

15 The base vector sequence may be linked to a transgene sequence that directs homologous recombination of the vector into the genome of a host cell. Such transgene sequences may or may not be from the same viral serotype as the base vector sequence.

The vector may be used to transport the heterologous sequence into a host cell or into a host cell's genome.

20 The vector may comprise multiple endonuclease restriction sites that enable convenient insertion of exogenous DNA sequences.

The term "hybrid vector" as used in the invention refers to a vector which comprises a nucleic acid sequence combined from two different viruses (e.g. Adenovirus and AAV).

25 "Chimeric vector" refers to a vector which contains nucleic acid sequences that are unnatural to the base vector (i.e. sequences not occurring naturally or sequences not in their natural background including heterologous sequences). A chimeric vector as used in the invention may also be a hybrid vector. An example of a chimeric vector is Ad.AAV expressing a modified fiber protein on its capsid.

30

The term "transduction" or "infection" refers to a method of introducing viral DNA within a virus particle into a host cell. The viral DNA herein is in the form of recombinant virus, which is generated by linking a segment of DNA of interest into the viral genome in such a way that the gene can be expressed as a functional protein.

5

The term "transfection" refers to a method of introducing a DNA fragment into a host cell.

The term "heterologous" as used herein means that a nucleic acid or peptide sequence is placed in a context that is not endogenous to the base adenovirus vector or to a transduced cell. For example, a peptide sequence can be transferred from a protein to another protein, the resulting protein is referred to herein as heterologous protein. A chimeric fiber protein, (e.g., a serotype 5 tail domain and a serotype 35 shaft and knob domain) is considered a "heterologous" to the Ad 5 vector. The term also includes nucleic acids (e.g. coding sequences) from one strain or serotype of adenovirus introduced into a different strain or serotype of adenovirus.

The term "regulatory elements" is intended to include promoters, enhancers, transcription termination signals, polyadenylation sequences, and other expression control sequences. Regulatory elements referred to in the invention include but are not limited to, those which direct expression of nucleic acid sequence only in certain host cells (e.g. tissue specific regulatory sequences).

The term "operably linked" indicates that a polynucleotide sequence (e.g. a coding sequence or gene) is linked to a regulatory element in such a way that the regulatory element sequence controls and regulates the transcription or translation or both of that polynucleotide sequence. The orientation of the regulatory element may vary (eg, be in reverse orientation with respect to the right ITR). The term also includes having an appropriate start signal (e.g., ATG) in front of the polynucleotide sequence to be expressed and maintaining the correct reading frame to permit expression of the polynucleotide sequence under the control of the expression control sequence and

production of the desired polypeptide or protein. Regulatory sequences can also include 3' sequences which ensure correct termination (eg. polyadenylation stop signal).

5 The term "gene therapy" used herein, refers to a method which introduces a segment of exogenous nucleic acid into a cell in such a way that it results in functional modification to the recipient cell by expression of the exogenous nucleic acid. The exogenous nucleic acid is typically therapeutic in that the expression of the encoded protein, polypeptide or RNA corrects cellular dysfunction due to a genetic error or more generally counteracts any undesirable functions which are associated with a genetic or acquired disease. The  
10 term "exogenous nucleic acid" refers to DNA or RNA sequences not normally expressed in the treated transformed cell. The term also refers to DNA and RNA sequences which are expressed in a treated transformed cell at a higher, lower or in an otherwise different pattern than in the untreated, nontransformed cell. This non-natural expression can also be termed heterologous expression.

15 A "gene therapy vector" refers to a vector used for gene therapy. i.e. to introduce the exogenous nucleic acid into a recipient or host cell. The exogenous nucleic acid may be transiently expressed or integrated and stably expressed in the recipient or host cell.

20 The term "plasmid" as used herein refers to any nucleic acid molecule which replicates independently of the host, maintains a high copy number, and which can be used as a cloning tool.

The term "parallel strand of DNA" and "anti-parallel strand of DNA" refers to as each of  
25 the strands of DNA of the double stranded adenovirus. The Figures diagram the location of certain nucleotides on the parallel strand of DNA. The anti-parallel strand of DNA refers to the other of the two strands of DNA which is not depicted in the Figures. The fiber protein is encoded on the anti-parallel strand of DNA. To simplify the vector diagrams, the fiber sequences are shown on the parallel strand even though the gene is  
30 located on the anti-parallel strand.

The term "reporter gene" refers to any nucleic acid sequence which encodes a polypeptide or protein which can be readily detected by, for example, a visual, microscopic, immunochemical or enzymatic assay. Preferred reporter genes are those  
5 that can be detected by a non-destructive method that does not destroy the treated, transformed cells or tissue.

The term "selection gene" used herein refers to any nucleic acid fragment which encodes a polypeptide or protein whose expression is used to mark a cell as a transformed cell by  
10 a given vector.

The term "therapeutic gene" refers herein to a DNA fragment encoding a functional polypeptide, protein or RNA, which when expressed in a host cell provides a therapeutic benefit or desired function to the host cell or to the organ or organism containing the host  
15 cell. The therapeutic benefit may result from modification of a function of a native gene in a host or from the additional function provided by the therapeutic protein, polypeptide or RNA.

The term "host tissue" or "host cell" as used herein, refers to a tissue or cell in which a  
20 therapeutic gene is to be expressed to modify its function.

It is well-known in the biological arts that certain amino acid substitutions may be made in protein sequences without affecting the function of the protein. Generally, conservative amino acid substitutions or substitutions of similar amino acids are tolerated  
25 without affecting protein function. Similar amino acids can be those that are similar in size and/or charge properties, for example, aspartate and glutamate, and isoleucine and valine, are both pairs of similar amino acids. Similarity between amino acid pairs has been assessed in the art in a number of ways. For example, Dayhoff et al. (1978) in *Atlas of Protein Sequence and Structure*, Volume 5, Supplement 3, Chapter 22, pp. 345-352,  
30 which is incorporated by reference herein, provides frequency tables for amino acid substitutions which can be employed as a measure of amino acid similarity. Dayhoff et

al's frequency tables are based on comparisons of amino acid sequences for proteins having the same function from a variety of evolutionarily different sources. Therefore, any obvious changes in the amino acid sequences (as described above) to the sequences of the invention are already contemplated.

5

Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of  
10 amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; a group of amino acids having sulfur-containing side chains is cysteine and methionine.  
15 Preferred conservative amino acids substitution groups include but are not limited to: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, asparagine-glutamine, and aspartate-glutamate. Therefore, polypeptide substitution for "substantially similar" sequences (as described above) to the amino acid sequences described invention are already contemplated.

20

In order that the invention herein described may be more fully understood, the following description is set forth.

The present invention provides unique gene transfer vehicles which overcome many of  
25 the limitations of prior art vectors. The invention describes a first generation adenovirus vectors comprising left and right Ad ITRs, an Ad packaging sequence, a transgene cassette with regulatory elements, and a pair of cassette ITRs flanking the transgene cassette that direct predictable viral genomic rearrangements during viral replication as well as direct the integration of the transgene cassette into the host cell genome.

30

One predictable rearrangement that occurs during viral replication is the generation of a gutless adenovirus vector (also referred to herein as  $\Delta$ Ad) that comprises right and left Ad ITRs, an Ad packaging sequence, a transgene cassette flanked by cassette ITRs and the gutless vector is devoid of all other immunogenic viral genes.

5

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the transgene cassette is directed by co-infection with an Ad vector expressing e.g., the rep 78 protein to achieve site-specific integration in the e.g., AAVS1 site on human chromosome 19.

10

The invention further describes a novel way of targeting these recombinant adenovirus vectors to selected cells by modifying the adenovirus fiber protein that is expressed on the capsid. Changes to both the fiber shaft and the fiber knob domain proved to successfully retarget the Ad vector to a desired cell type. In addition, the G-H loop within the fiber knob domain is identified as a novel site that affect the binding affinity and specificity of the recombinant adenovirus vector. Substitution of peptide sequences into the G-H loop retarget the gutless vector to a desired cell type.

An adenovirus display library has been generated that expresses random peptides within the G-H loop of the fiber protein. This type of a library is used as ligands to screen for adenovirus vectors that bind to desired cell types. One advantage of using an adenovirus display library versus a phage display library is that once adenovirus affinity to a desired cell is identified the targeted adenovirus vector is ready to accept a transgene cassette and can be used to generate a gutless adenovirus vector, for example for use in gene therapy.

25

The chimeric Ad vectors described below contain a modified fiber protein in the capsid of the adenovirus which renders the vector capable of infecting a desired cell types. Therefore, according to the invention, a gutless chimeric  $\Delta$ Ad-AAV vector can be generated to introduce any transgene(s) into any host cell or tissue which is normally refractory to most commonly used gene therapy viral vectors. In addition, the chimeric  $\Delta$ Ad-AAV vector of the invention, is devoid of adenoviral genes, and contains AAV ITR

30

sequences that flank the transgene cassette, which direct stable transgene integration in the host genome allowing long term expression of the transgene.

The transgene cassette described in the invention may carry a transgene which is either a reporter gene, a selectable gene for in vitro or in vivo selection of transduced cells, or a therapeutic gene. In one embodiment of the invention the reporter transgene can be but is not limited to,  $\beta$ galactosidase. Many reporter genes are commonly used in the art, of which any could be carried as a transgene in the Ad.AAV vector of the invention. Other examples of reporter genes are GFP and alkaline phosphatase.

The following describes an embodiment of the first generation Ad vectors of the invention having a wild-type capsid and a transgene cassette flanked by cassette ITR sequences; (b) fiber protein that is modified to retarget Ad vectors; and (c) the combination of both technologies that enables the production of chimeric  $\Delta$ Ad vector including a modified fiber protein expressed on the capsid which retargets the base vector to a desired cell type for infection and transgene integration.

A. Integrating Ad hybrid vectors of the invention:

It has been shown that inverted repeats (IRs) inserted into the E1 region of AdE1- vectors can mediate predictable genomic rearrangements resulting in a gutless vector genome devoid of all viral genes. A specific embodiment of such IR-mediated rearrangements is the Adeno-AAV, first generation hybrid adenovirus vector containing AAV inverted terminal repeats (ITR) flanking a transgene cassette. The AAV ITRs mediate the formation of a genome similar to that of the  $\Delta$ Ad.IR genome (Steinwaerder et al., 2000 Journal of Virology).  $\Delta$ Ad vectors devoid of all viral genes stably integrate and transduce cultured cells with efficiencies comparable to e.g. rAAV vectors. The Examples demonstrate by Southern blot analysis that the  $\Delta$ Ad vectors integrate randomly into the host genome.

The Ad vectors of the invention comprise a left Ad ITR, an adenovirus packaging sequence located 3' to the Ad ITR; a transgene cassette located 3' to the packaging

sequence comprising a polyadenylation signal, a transgene, and a heterologous promoter, and flanked by a pair of cassette ITRs. Adenoviral genes used for replication such as E1, E2, E3, E4 are located 3' to the right cassette ITR and a right Ad ITR is located 3' to the replication genes. The vectors of the invention are particularly suited to treat: genetic disorders, cancers, and infectious diseases (such as HIV, emboli, or malaria). Treatable genetic diseases such as hemophilia A and B; cystic fibrosis; muscular dystrophy, and  $\alpha_1$ -antitrypsin disorder are ideal candidates for genetic disease that can be treated by vectors of the invention. A specific example of a therapeutic gene to combat a genetic disorder is gamma-globin to ameliorate sickle cell anemia.

10

To aid in the selection of transduced cells and characterize the integration site of the transgene cassette, an embodiment of the invention includes the addition of a sequence comprising a bacterial origin of replication, plus a selectable gene. An embodiment of this is an SNori sequence added to the transgene cassette. This allows the  $\Delta$ Ad to be expressed in human and bacterial cells, therefore allowing selection of the transduced cells and characterization of the integration site in the genome of transduced mammalian cells.

15

The potential for site-specific integration is an important characteristic of the novel Ad vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein in 293 cells to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration, which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

25  
30



Integration of the transgene cassette contained in the adenoviral vectors into chromosomes may be associated with silencing (or blocking) of transgene expression. The silencing of transgenes can be overcome by adding insulator elements to the transgene cassette. For example, HS-4 insulator elements derived from the chicken-globin LCR can function in Ad vectors to shield heterologous promoters from adenoviral enhancers. HS-4 insulators or the *Drosophila* Gypsy gene can also be used to prevent silencing transgenes.

Another embodiment of the invention is to split the transgene cassette into two portions of the transgene each carried in a different recombinant adenoviral vector of the invention. Each portion of the same transgene has an overlapping region of homology. After infection with both vectors, each carrying the different but overlapping portion of the same transgene, homologous recombination event occurs resulting in the reconstitution of the complete transgene which is then expressed. This technique is used to produce hybrid adenoviral vectors that accommodate large inserts including, but not limited to a 13kb genomic hAAT gene or a 12kb  $\gamma$ -globin LCR $\gamma$ -globin expression cassettes for ameliorating sickle cell anemia (or correcting  $\gamma$ -globin mutations). The formation of the hybrid  $\Delta$ Ad vector genomes, after recombination between two vectors, is more efficient if the overlapping region of homology within the transgene is longer.

An advantage of the present invention is a method to rapidly isolate pure gutless hybrid adenoviral vectors such as  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV<sup>fx</sup> vectors. To minimize the contamination of  $\Delta$ Ad with first generation vectors (Ad vectors) a strategy is described in Example I H. It is anticipated that these approaches will yield the same titer of  $\Delta$ Ad vectors, however the contamination with full-length genome vectors will be less. This improved isolation of the vectors is extremely important to avoid toxic side effects after in vivo application.

**B. Tropism modified adenovirus vectors:**

The Ad vectors of the invention can be modified so that they target a host cell of interest. There are more than 50 human Ad serotypes (Appendix I), including variants with different tissue selectivity or tropism. It is accepted in the art that different Ad serotypes bind to different cellular receptors and use different entry mechanisms. Most recombinant  
5 adenovirus vectors use adenovirus serotype 5 as the base vector serotype 5 (Ad5) (Hitt, M.M., et al, 1997, Adv. in Pharmacology 40, 137-205). Ad5 infection is primarily mediated by its fiber protein binding to CAR and secondarily by its penton base protein binding to integrin. Due to the lack of CAR and/or integrin expression on many cell and tissue types, Ad5 mediated gene transfer is inefficient in a number of tissues which are  
10 important targets for gene therapy such as endothelia, smooth muscle, skin epithelia, differentiated airway epithelia, brain tissue, peripheral blood cells, or bone marrow. The following describes Ad5 vectors of the invention having a change in infectivity and tropism as a result of altering the fiber protein sequence.

15 The infectivity of different Ad serotypes is limited to a number of human cell lines. Infectivity studies revealed that Ad5 and Ad3 are particularly suitable for infecting and targeting endothelial or lymphoid cells, whereas Ad9, Ad11 and Ad35 efficiently infected human bone marrow cells. Therefore, the knob domain of the fiber protein of Ad9, Ad11 and Ad35 are excellent candidates for retargeting the Ad5 vector to human bone marrow  
20 cells. Other possible serotypes include Ad7.

In the modified fiber protein of the invention the fiber knob domain of the Ad5 fiber has been replaced with another Ad serotype fiber knob domain. An embodiment of the invention is the modified Ad5/35 fiber protein (a recombinant Ad5 vector expressing a  
25 modified fiber protein comprising of a fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad35). The Ad5/35 chimeric fiber protein shows a broader spectrum of infection to a subset of CD34+ cells, including those with stem cell activity. The Ad5/11 chimeric fiber protein (a recombinant Ad5 vector expressing a modified fiber protein comprising the fiber tail domain of Ad5 and the fiber shaft and knob domains of Ad11)  
30 showed similar tropism.

In addition to the knob domain modifications, the invention describes the added advantage of modifying both the fiber shaft domain and the fiber knob domain to produce a shortened fiber protein. The length of the fiber shaft domain plays a key role in the host receptors used for viral vector entry into the host cell. To show this Ad5, Ad5/9, and  
5 Ad5/35 variants were constructed with long (22 $\beta$ -sheets) and short-shafted (7 $\beta$ -sheets)-shafted fibers. These analyses demonstrated that efficient viral infection involving CAR as the primary receptor for Ad5, Ad5/9 requires a long-shafted fiber protein, whereas the cell entry strategy of Ad5/35 (which binds to an still uncharacterized non-CAR receptor) does not depend on the shaft length (Fig. 3). The modification in both the fiber shaft  
10 domain length (between 5 > 10  $\beta$ -sheets) and the fiber knob domain (from a different Adserotype than the base vector is a novel mode of altering Ad vector tropism.

To broaden the repertoire of cell types that Ad vectors can infect, a specific binding region, the G-H loop, within the knob domain has been newly identified herein to  
15 improve binding affinity and specificity. Alteration within this region will redirect the Ad vector to a desired cell type. For example, the invention describes the G-H loop sequence within the fiber protein knob domain, which can be replaced with heterologous peptide ligand sequences without affecting the functionally important tertiary structure of the Ad fiber knob domain, while changing the binding affinity and specificity of the vector (Figs.  
20 6,7). This G-H loop region is exposed on the central part of the knob surface and may be strategically a better site for incorporation of heterologous ligands than the peripheral H-I loop (Krasnykh, V. et al., 1998, *J. Virol.*, 72:1844-52.) of the knob C-terminus (Michael, S. I., et al., 1995, *Gene Ther.*, 2:660-8., Wickham, T. J. et al., 1996, *Nat. Biotechnol.*, 14:1570-3.), which are the substitution sites used by others. Therefore, these G-H loop  
25 modifications within the fiber knob domain will allow the Ad vector to be redirected to infect a desired cell type, as long as the G-H loop ligand sequence binds to at least one surface protein on the desired cell type. Fig. 7 shows some possible substitutions. Example II J demonstrates that the virion tolerates the insertion of a cycling peptide (12 amino acids) with a constrained secondary structure that allows the exposure on the knob  
30 surface. A defined ligand (RGD) can be inserted into the G-H and the H-I loop of an

Ad5 capsid that is ablated for CAR, and integrin tropism. Infectivity studies show the potential advantage of this new insertion site.

Use of the Vectors of the Invention for "Gene Therapy"

5

The liver is the major organ for protein synthesis. Therefore an important goal of gene therapy is to target gene therapy vectors to the liver. To genetically correct many types of mutant proteins hepatocytes need to be infected with gene therapy vectors carrying a corrected transgene. Example II J describes a G-H loop substitution in the knob domain of the fiber protein with both RI and RII+ (of the malaria circumsporozoite surface protein) in a short shafted fiber protein which directs the vector to have affinity and specificity to hepatocytes.

10

Example II K applies a similar protocol to alter the fiber knob domain in the G-H loop region with peptides that target the vector breast cancer cell lines (MDA-MB-435). These novel approaches to redirect vectors described in the invention allow lower doses of the gene therapy vectors to be administered with a higher safety profile.

15

In example II L a protocol for preparing an adenovirus display library is described that uses the fiber knob protein to display a library of random peptide sequences within the G-H loop. This library of adenoviruses with modified fiber proteins is screened for affinity and specificity for a desired cell type. There are two main advantages of using this adenovirus display library to screen for target peptides that allow binding to a desired cell type over a phage display library system. First, once a ligand peptide is identified that binds to the desired cell type it is already in the vector of choice for gene therapy delivery. The peptide does not need to be engineered into another vector, as is the case for the phage display library vectors. This reduces the steps required to identify a targeted fiber protein for a desired cell type. The second advantage of this method, is that the adenoviruses are able to display multiple copies of the modified fiber protein on their capsid. This allows for dimerization and trimerization of the fiber protein with the host cell receptor. The multimerization of fibers proteins is a realistic, in vivo interaction of

20

25

30

the trimeric fiber protein with the host cell receptor. In contrast, phage vectors can only display one fiber peptide sequence on their surface, which significantly limits the ability of interaction with host cell surface receptors.

5    C. A chimeric adenovirus vector with selective tropism:

The chimeric vectors of the invention combine two vectors: an Ad.ITR and a Ad. fx where fx describes a modified fiber protein. A first generation adenovirus vector of serotype 5 is the base vector that carries a transgene cassette flanked by heterologous  
10    ITRs. These specific inverted terminal repeat sequences, such as AAV ITRs direct stable integration of the transgene cassette into the host genome as well as control predictable genomic rearrangements that occur during viral replication. This vector can also carry a modified fiber gene (described in Examples II). During replication predictable genomic rearrangements occur which result in the generation of a gutless adenovirus vector (e.g.  
15     $\Delta$ Ad.AAV<sup>fx</sup>) which expresses the modified fiber protein on its capsid. The modified fiber protein allows the gutless vector to be targeted to a selected cell type. The targeted vector is a gutless adenovirus vector devoid of adenoviral genes which can integrate its transgene into the host genome. The transgene cassette can carry reporter, selectable, or therapeutic genes.

20

In one embodiment of the invention, the gutless targeted  $\Delta$ Ad.AAV<sup>fx</sup> carries the reporter gene of  $\beta$ galactosidase ( $\Delta$ Ad.AAV<sup>fx</sup>-BG). For easy in vitro selection of human and bacterial cells that are transduced with the hybrid Ad vector, a bacterial sequence for the origin of replication can be added to the hybrid Ad vectors. An example of this is  
25     $\Delta$ Ad.AAV<sup>fx</sup>-Snori, in which a SNori sequence is added into the transgene cassette. This site allows for G418 selection on cells infected with  $\Delta$ Ad.AAV<sup>fx</sup>-SNori. This in vitro selection provides a tool to analyze the site of transgene integration and the flanking chromosomal regions. Fluorescent in situ hybridization (FISH) is an alternative method to confirm vector integration.

30

An advantage of the  $\Delta$ Ad chimeric vector for gene transfer is the efficient and stable integration of a large transgene cassette up to e.g. 22kb which is significantly larger than the capacity of retroviral vectors. This is of particular interest for gene therapy. For example, to ameliorate sickle cell anemia  $\Delta$ Ad.AAV<sup>fx</sup>  $\gamma$ globin, an expression transgene cassette with the gamma-globin gene that targets and integrates, can be inserted into bone marrow stem cells for long term expression of the gamma-globin gene.

To achieve site-specific gene integration, rep78 protein is used for transgene integration into the AAVS1 site (described in Example 1 D). However, this may silence transgene expression. To prevent the integrated transgene from being silenced by host genomic elements (such as positional effects or downstream enhancers), LCRs or insulator elements are incorporated into the transgene cassette.

The following examples are presented to illustrate the present invention and to assist one of ordinary skill in making and using the same. The examples are not intended in any way to otherwise limit the scope of the invention.

#### **EXAMPLE I**

##### **NOVEL ADENOVIRAL VECTOR Ad.AAV**

##### **A. Integrating $\Delta$ Ad.AAV hybrid vectors devoid of all adenoviral genes.**

*In vitro* and *in vivo* studies with rAAV indicate that the only requirement for rAAV integration are the AAV ITRs and as yet unknown host cellular factors. It is thought that specific sequences or secondary structures present in AAV ITRs are prone to integration into host chromosomal DNA. In order to combine advantages of adenoviral vectors (high titer, high infectivity, large capacity) and the integration capability of AAV ITRs, AAV vector DNA with AAV ITRs flanking cassettes a secreted human placental alkaline phosphatase (SEAP) - neomycin phosphotransferase (neo) reporter gene cassette

(Alexander, I.E., et al. 1996, *Gene Therapy*, 7, 841-850) is incorporated into the E1-region of E1/E3 deleted adenoviral vectors (Ad.AAV1) (Figure 1, top).

## METHODS

5

### Production/Characterization of Viral Vectors

#### Plasmids:

- 10 The AAV1 vector cassette containing AAV ITRs and SEAP/neo expression units is obtained by AseI/ScaI digestion of the plasmid pALSAPSN (Alexander, I. E. et al, 1996, *Human Gene Therapy*, 7:841-50). The 4.4kb AAV vector fragment was cloned via NotI adapter linkers into pXJCL1 (Mirobix, Toronto, Canada) (pAd.AAV1). Another shuttle vector (pAd.AAV1- $\Delta$ 2ITRs) lacking the AAV ITRs is generated by inserting the 3.7kb
- 15 AflII/BsmI fragment of pALAPSN into pXJCL1. For pAd.AAV1 $\Delta$ 1ITR, a construct is used where a spontaneous deletion in the left AAV ITRs between the A and A' regions has occurred. To create a second hybrid vector (Ad.AAV2), the AAVSNori cassette developed by E. Rutledge is used. AAV vector DNA obtained is from pASNori (Rutledge, E. A., Russell, D.W. 1997. *Journal of Virology* 71:8429-8436) as a 3.4kb
- 20 BsaI/ScaI fragment and inserted into the EcoRV site of pXCJL1. As it is generally known for AAV vector plasmids, the AAV ITRs are prone to rearrangements. To minimize deletions in these functional critical regions, all constructs for generation of hybrid vectors are assembled in low copy-number plasmids which are grown in *E. coli* Top10, JC811, or XL1 Bluecells (Stragene, La Jolla, Calif.). Furthermore, after each
- 25 cloning step or large-scale plasmid amplification, both AAV ITRs are carefully mapped by restriction analysis with enzymes that cut inside or adjacent to the ITRs (BssHII, AhdI, SmaI, BglI, BsmI, AflII, and ScaI).

Adenoviruses:

First-generation viruses with the different transgene cassettes incorporated into the E1 region are generated by recombination of the pΔE1aSplα- or pXCJL1-derived shuttle  
5 plasmids and pJM17 (Microbix) in 293 cells as described earlier (Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791). For each virus, at least 20 plaques are picked, amplified, and analyzed by restriction digest. Viruses containing two AAV ITDRs tend to rearrange within the ITRs, with other adenoviral sequences, or with adenoviral sequences present in the 293 cell genome. Only plaques from viruses with intact ITRs are amplified, CsCl  
10 banded, and titered as described earlier (Kay, M. A., et al. 1995. *Hepatology* 21:815-819; Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960). All virus preparations tested are negative for RCA and bacterial endotoxin (Lieber, A., et al. 1997. *Journal of Virology* 71:8798-8807). Virus is stored at -80°C in 10mM Tris-Cl, pH 7.5-1 mM MgCl<sub>2</sub>-10% glycerol.

15 To generate ΔAd.AAV, 293 cells are infected with Ad.AAV1 at an multiplicity of infection (MOI) of 25 and harvested 40h after infection. Cells are lysed in PBS by 4 cycles of freeze/thawing. Lysates are centrifuged to remove cell debris and digested for 30 min at 37°C with 500 units/ml DNaseI and 200 μg/ml RNaseA in the presence of  
20 10mM MgCl<sub>2</sub>. 5ml of lysate is layered on a CsCl step gradient (0.5ml - 1.5g/cm<sup>3</sup>, 2.5ml - 1.35g/cm<sup>3</sup>, 4ml - 1.25g/cm<sup>3</sup>) and ultracentrifuged for 2h at 35,000 rpm (rotor SW41). CsCl fractions are collected by puncturing the tube and are analyzed for viral DNA (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960; Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) or subjected to ultracentrifugation at 35,000 rpm for 18 hours  
25 in an equilibrium gradient with 1.32 g/cm<sup>3</sup> CsCl. The band containing the deleted viruses ΔAd.AAV is clearly separated (0.5cm distance) from other banded viral particles containing full-length ad.AAV genomes. ΔAd.AAV1 fractions are dialyzed against 10mM Tris-Cl, pH 7.5-1mM MgCl<sub>2</sub>-10% glycerol and stored at -80°C. The genome titer of ΔAd.AAV1 preparations is determined based on quantitative Southern analysis of viral  
30 DNA purified from viral particles in comparison to different concentrations of a 4.4kb AseI/ScaI fragment of pALSAPSN according to a protocol described earlier (Lieber, A.,



et al., 1996, *J. of Virology*, 70, 8782-8791). In total, the production of  $1 \times 10^{13}$  genome-particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of actual work.

5 Titers routinely obtained are in the range of  $3-8 \times 10^{12}$  genomes per ml. Assuming one genome is packaged per capsid, the genome titer equals the particle titer. The level of contaminating Ad.AAV1 is less than 0.1% as determined by Southern analysis, which is consistent with results obtained by plaque assay on 293 cells (fewer than 5 plaques per  $10^6$  total genomes). The primers used for sequencing the left and right ITR-vector-junction are

10

5'GGCGTTACTTAAGCTAGAGCTTATCTG, and  
5'CTCTCTAGTTCTAGCCTCGATCTCAC.

The recombinant AAV virus stock containing the SEAP/neo cassette (AV2/ALSAPSN,  
15 [Alexander, I. E. et al, 1996, *Human Gene Therapy*, 7:841-50] used in these studies were obtained from Dusty Miller (FHCRC, Seattle). The stock was free of contaminating replication competent AAV ( $<50$  particles/ml) and wildtype adenovirus ( $<100$  particles/ml). The genome titer of the virus stock was obtained by quantitative Southern Blots as described by Russell et al. (Russell, D. et al. 1994 *Proc. Natl.Acad.Sci. USA*  
20 91:8915-8919).

#### Electron Microscopy:

For examination of viral particles in the transmission electron microscopy studies, CsCl-  
25 purified virions are fixed with glutaraldehyde and stained with uranyl acetate as described previously (Lieber, A., et al. 1996. *Journal of Virology* 70:8944-8960).

#### RESULTS

30 During replication of these hybrid vectors in 293 cells, a 5.5kb genome ( $\Delta$ Ad.AAV1) is efficiently generated and packaged into adenovirus (Ad5) capsids. The  $\Delta$ Ad.AAV1

genome contains the left adenovirus ITR and the packaging signal followed by the AAV-vector cassette and a duplicate of the adenoviral packaging signal and ITR in reverse orientation (Figure 1, bottom). The hybrid vector is devoid of all viral genes, thus eliminating toxic effects and the elicitation of cellular immune responses. The spontaneous formation of the small hybrid vector genome  $\Delta$ Ad.AAV1 requires the presence of two intact AAV ITRs and does not occur with partly deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette.

Hybrid vectors containing different transgenes:

To construct a hybrid vector with a transgene that can be detected in situ the SEAP/neo expression unit in Ad.AAV1 is replaced by the *E. coli*  $\beta$ -galactosidase gene. This hybrid vector is named  $\Delta$ Ad.AAV1. During generation of the corresponding plasmid constructs the AAV ITR sequences tend to rearrange and abolish their functional properties. This problem can be circumvented by using low copy number plasmids as cloning vectors grown in bacteria strains depleted for all recombination proteins (e.g.JC811). Furthermore, the intactness of both AAV ITRs after each cloning step can be examined for characteristic endonuclease digestion. Recently, another hybrid vector  $\Delta$ Ad.AAV1Nori has been generated which contains the neo gene under the control of both the simian virus 40 (SV40) early promoter and the transposon 5 (Tn5) promoter for expression in human and bacterial cells, as well as the p15A bacterial replication origin with the direction of the leading strand DNA synthesis opposite that of neo gene transcription. Thus, SNori can be used for G418 selection of integrated vector in eukaryotic cells as well as for rescue of vector together with flanking host DNA after integration. The recovered plasmids can be propagated in *E. coli* under selection with kanamycin due to the bacterial origin and the neo gene. SNori containing vectors allow a rapid estimation of total integration events based on the number of G418 resistant colonies. Moreover, vector DNA together with flanking chromosomal DNA can be rescued as plasmids from single G418 resistant clones and can be used for sequencing to determine integration junctions. Both hybrid vectors are produced at a titer of about

$3 \times 10^{12}$  genomes per ml. The ratio of genome titer to transducing particles for  $\Delta$ Ad.AAVBG is  $\sim 200:1$  based on  $\beta$ -Gal expression.

## DISCUSSION

5

$\Delta$ Ad.AAV1 could spontaneously form during adenovirus replication. Another possible mechanism of  $\Delta$ Ad.AAV1 formation is based on the unique mechanism by which adenovirus replicates its genome (van der Vliet, B., 1995, *In w. Doerfler, et al. (eds.) vol. 2 p. 1-31, Springer-Verlag, Berlin*) (see Figure 1). Ad DNA replication is initiated by the TP/pTP (terminal protein) that binds to specific sites within the ITRs on both ends of the linear genome and catalyzes, in complex with Ad pol, the binding of the 5' CTP, the first nucleotide of the daughter strand. DNA synthesis proceeds in a continuous fashion to the other end of the genome (Figure 1A). Only one of the DNA strands serves as template. One of the replication products is a single-stranded DNA that circularizes through annealing of its self-complementary ITRs. The resulting duplex "panhandle" has the same structure as the termini of the duplex viral genome that allows the binding of pTP and the initiation for synthesis of a complementary strand using the single-stranded "pandhandle" molecule as template (Figure 1C). In the case of Ad.AAV1, the Ad pol synthesizes the single strand of the adenoviral genome starting from the left Ad ITR until it reaches the second AAV ITR. During synthesis of the second AAV ITR a certain percentage of the single-stranded molecules form a loop hybridizing to the complementary region within the first AAV ITR that was replicated earlier, allowing Ad pol to use the same viral DNA strand to read back towards the left ITR (Figure 1B). The resulting "panhandle" structure can be resolved in a similar way as a full-length intermediate shown in Figure 1C, generating a double stranded, linear molecule with the above described structure that can be packaged into Ad virions. The ratio of viral DNA to protein concentration in purified  $\Delta$ Ad.AAV1 particles is comparable to that obtained from Ad.AAV1 particles. This indicates that despite the smaller size, only one  $\Delta$ Ad.AAV1 genome is packaged, resulting in particles with a lighter buoyant density ( $\sim 1.32 \text{ g/cm}^3$ ). Electron microscopy demonstrates the icosahedral shape of  $\Delta$ Ad.AAV1 particles (Figure 2). Staining with uranyl acetate causes the central viral cores to appear

30

electron dense.  $\Delta$ Ad.AAV1 virions have only a spotted luminal dark staining as expected with only one 5.5kb genome being packaged per capsid.

**B. In vitro  $\Delta$ Ad.AAV1 production:**

5

**Characteristics of deleted adeno-AAV vectors ( $\Delta$ Ad.AAV):**

A number of experiments to clarify the mechanisms of  $\Delta$ Ad.AAV genome formation are carried out. Specifically, the presence of two intact AAV ITRs flanking a reporter gene cassette is required for the effective formation of  $\Delta$ Ad.AAV genomes. This process does not occur with partially deleted ITRs or oligo-dC and oligo-dG stretches flanking the expression cassette. Furthermore, in vitro transduction studies are performed with different genome titers of  $\Delta$ Ad.AAV1, Ad.AAV1, and Ad.AAV1- $\Delta$ 2ITRs, (lacking the two AAV ITRs) which determine the number of G418 resistant colonies that formed after 4 weeks of selection (Table I).

$\Delta$ Ad.AAV1 is routinely produced at a high titer ( $5 \times 10^{12}$  genomes per ml with  $>10^4$  produced genomes per 293 cell) and at a high purity with less than 0.1% contaminating full length Ad.AAV1 genomes by a technique normally used for amplification and purification of recombinant adenovirus.

**In vitro transduction studies with hybrid vectors on CD34+ cells and erythroleukemia cells:**

In order to test whether the hybrid vectors allow for gene transfer into cell types, that have to be targeted for sickle cell therapy, infection/transduction studies are performed using CD34+ enriched human bone marrow cells, derived from mobilized peripheral blood and the human erythroleukemia cell line K562 which express  $\epsilon$  and  $\gamma$  globin genes.

30 **METHODS**

Cell Culture:

SKHep1 cells (HTB-52, American Type Culture Collection, Rockville, MD), an endothelial cell line derived from human liver [Heffelfinger, S.C., et al., 1992, *In vitro Cell Dev. Biol.* 28A, 136-4-142], are grown in high-glucose Dulbecco's modified Eagle medium with 10% fetal calf serum. SKHep1 cells are analyzed for integrated AAV provirus by Southern analysis of genomic DNA using the AAV1 wild type genome obtained from pAAV/Ad (Samulski, R. J., et al. 1989. *Journal of Virology* 63:3822-3928) (gift from David Russell, University of Washington) as a probe. No specific bands are detected in undigested genomic SKHep1 DNA or after digestion with HindIII. For viral infection, confluent cells are incubated with different viral doses for 2 hours, followed by intensive washing. For G418 selection, 24h after infection with  $\Delta$ Ad.AAV1, SKHep1 cells are trypsinized and plated at different dilutions under G418 selection (900 $\mu$ g/ml active compound, Boehringer-Mannheim, Germany). G418 containing culture medium is changed every 3 days. The number of colonies with  $>10^4$  cells is counted after 4 weeks of selection and divided by the number of initially seeded cells. This ratio is used to express the integration frequency of  $\Delta$ Ad.AAV1. Single colonies are obtained by limiting dilutions of infected cells in 96 well plates. Colonies are expanded to  $1 \times 10^6$  cells in the presence of G418. Immunofluorescence analysis for adenoviral proteins expressed in SKHep1 cells 3 days post-infection is performed as described earlier [Lieber, A., et al., 1996, *J. of Virology*, 70, 8782-8791].

RESULTS

293 cells are infected with the first generation vector Ad.AAV1. During replication of Ad.AAV1, the small  $\Delta$ Ad.AAV1 genome forms spontaneously and is packaged into adenovirus capsid. At 36 hours after infection cells are harvested and virus is released by several cycles of freeze/thawing. The mixture of Ad.AAV1 and  $\Delta$ Ad.AAV1 particles in the cell lysate is then separated by ultracentrifugation in a CsCl step gradient. Due to its lighter buoyant density, the band containing the  $\Delta$ Ad.AAV1 particles is clearly separated (0.8cm distance) from the band containing full-length virus (Lieber, A., et al. 1999. *J*

*Virol* 73:9314-24).  $\Delta$ Ad.AAV1 is purified further by an additional CsCl equilibrium gradient and is stored in 10mM Tris pH7.5, 10% glycerol, 1mM MgCl<sub>2</sub> in 80°C. In total, the production of  $2 \times 10^{13}$  (genome) particles of  $\Delta$ Ad.AAV1 requires less than 3 hours of work. All functions for  $\Delta$ Ad.AAV1 replication and particle formation are provided from Ad.AAV1 genomes amplified in the same cell. The efficiency of vector production measured on a genome-per-cell-basis is comparable or higher than labor-intensive, newer techniques for rAAV production, which have not yet been proven to be reliable. The estimated ratio of transducing/genome titer for  $\Delta$ Ad.AAV1 is 1:200 (based on SEAP expression at day 3 post-infection), whereas for the average rAAV preparation, it is in the range of 1:10<sup>3</sup> to 1:10<sup>4</sup>.  $1 \times 10^5$  confluent SKHepl cells are infected with different MOIs of rAAV1 (stock:  $1 \times 10^{10}$  genomes per ml),  $\Delta$ Ad.AAV1 (stock:  $5 \times 10^{12}$  genomes per ml), Ad.ADAV1 (stock:  $1 \times 10^{13}$  genomes per ml), and AdAAV1.2ITR (stock:  $9 \times 10^{12}$  genomes per ml), in a volume of 100ml 24 hours after infection, cells are washed, trypsinized, and plated at different dilutions. G418 is added 24 hours after plating and selection is performed for 4 weeks. G418 resistant colonies contain on average  $>5 \times 10^4$  cells (at least 16 cell divisions). A significant number of small colonies visible at 2 weeks post-infection do not survive continued selection, probably due to episomal vector expression. Cells infected with first-generation adenoviruses with MOIs greater than  $1 \times 10^4$  develop CPE during the first week of selection. The rAAV titer is not high enough to perform infection studies with MOIs greater than  $10^4$ . The colony formation is expressed as percentage of the number of colonies after selection to the number of cells initially seeded for selection (Table I).

**TABLE I**

Formation of G418 resistant colonies after infection with hybrid viruses in comparison with rAAV.

| MOI<br>(genomes per cell) | Formation of G418 resistant colonies in % (SEM)<br>(after 4 weeks of selection) |
|---------------------------|---|
|---------------------------|---|

|                 | rAAV1      | ΔAd.AAV1   | Ad.AAV1    | Ad.AAV1<br>ΔITRs |
|-----------------|------------|------------|------------|------------------|
| 10 <sup>1</sup> | 0          | 0          | 0          | 0                |
| 10 <sup>2</sup> | 0          | 0          | 0          | 0                |
| 10 <sup>3</sup> | 2.7 (1.6)  | 1.3 (1)    | 5.4 (3.0)  | 0                |
| 10 <sup>4</sup> | 90.8 (7.0) | 48.0 (8.9) | 12.9 (7.2) | 0                |
| 10 <sup>5</sup> | N/A        | 93.1 (5.4) | 3.8 (2.1)  | 0                |
| 10 <sup>6</sup> | N/A        | 100        | 0          | 0                |
| 10 <sup>7</sup> | N/A        | 100        | 0          | 0                |

N = 3 (SEM is indicated in parentheses.)

K562 cells are infected with different MOIs of ΔAd.AAVBG (1-10<sup>8</sup> genomes per cell):

- 5 Three days after infection, the total number of viable cells (based on Trypan blue staining) and the percentage of infected cells (based on X-Gal staining) are determined for all MOIs. The results are presented in Figure 4A.

Initial integration studies:

10

K562 cells are incubated with ΔAd.AAVSNori at an MOI of 2x10<sup>5</sup> genomes per cell and the colonies that formed after 4 weeks of G418 selection are counted in 96 well plates. G418 resistant colonies contain on average >5x10<sup>4</sup> cells which means that the original cell underwent at least 16 cell divisions.

15

Infection studies with ΔAd.AAVBG (1-10<sup>8</sup> genomes per cell) on CD34<sup>+</sup> cells:

- Cell infection on CD34<sup>+</sup> are as described for K562 cells. CD34<sup>+</sup> cells are cultured in IMDM supplemented with 20% FCS, kit ligand (stem cell factor-SCF) (100ng/ml), and  
 20 IL-3 (100ng/ml). Since a number of reports suggest that specific cytokines like GM-CSF or M-CSF which induce stem cell differentiation can stimulate integrin expression and may therefore affect internalization of Ad5 vectors, infection rates are compared with

Ad5 based hybrid vectors on CD34+ cells cultured with and without pre-stimulation with GM-CSF (50ng/ml) or M-CSF (50U/ml). The number of infected cells is counted based on X-Gal staining at day 3 after infection. To test for dose-dependent toxicity, viable cells are counted (based on trypan blue exclusion) at day 3 post-infection. Furthermore, whether high viral doses affect the ability of CD34+ cells to differentiate in methyl cellulose colony assays in presence of IL-3 and SCF is analyzed. The results are expressed as viable cells/X-Gal positive cells vs MOI (see Figure 4B).

## DISCUSSION

10

The above data demonstrate that  $\Delta$ Ad.AAV transduces stably an immortalized human cell line with a low frequency comparable to rAAV, however, transduction rates could be scaled up to 100% by using greater MOIs of  $\Delta$ Ad.AAV1, which is produced at higher titers than rAAV1. In contrast to infection with the first-generation vector, Ad.AAV1, infection with  $\Delta$ Ad.AAV1 is not associated with dose-dependent cytotoxicity because no viral proteins are expressed from these vectors in transduced cells. Furthermore, viral proteins present in the incoming  $\Delta$ Ad.AAV1 particles are not problematic in the dose range use. The comparison of transduction rates of  $\Delta$ Ad.AAV/Ad.AAV1 with the vector lacking AAV ITRs, Ad.AAV- $\Delta$ 2ITRs, supports the hypothesis that the presence of two intact AAV ITRs is crucial for hybrid vector integration.

20

The data demonstrate that the leukemia cell line can be infected at ~90% efficiency with Ad5 based hybrid vectors at MOIs  $2 \times 10^5$  genomes per cell without significant toxic side effects. However, this dose is still ~100 times greater than the dose necessary to infect 100% of HeLa cell, hepatoma cells, primary hepatocytes and other cell lines generally considered as permissive for Ad5 vector infection.

25

Since viral DNA in cells infected with  $2 \times 10^4$  genomes (or 100 transducing particles per cell) should be lost after 7 cell divisions, the presence of G418 resistant cells in the observed colonies suggests that  $\Delta$ Ad.AAVSNori genomes are integrated into or stably associated with the host genome. Based on the number of G418 resistant colonies one

30



out of 25,000  $\Delta$ Ad.AAVSNori genomes integrates stably into K562 cells. This is in agreement with the results obtained earlier with  $\Delta$ Ad.AAV1 in SK Hep1 cells.

5 The maximal dose used for infection of CD34+ cells ( $1 \times 10^8$ ) results in X-Gal staining of only ~10% of cells independently of GM-CSF/M-CSF. This demonstrates the obvious inability of Ad5 to infect CD34+ cells and is probably caused by the absence of specific receptors and/or integrins on the cell surface. CD34+ cells tolerate a large range of viral doses ( $1-10^7$ ) without obvious effects on cell viability and total cell number. This is not surprising because in order to develop toxic side effects adenovirus has to enter the cell and express viral genes. Hybrid vectors can be produced at titers of  $5 \times 10^{12}$  genomes per ml. Thus, the maximal MOI that can be used for infection (of  $10^4$  cells) is  $\sim 5 \times 10^8$  (in 100  $\mu$ l storage buffer). Based on the infection studies with  $\Delta$ Ad.AAVBG this dose may not be sufficient to efficiently transduce CD34+ cells and to obtain an appreciable number of G418 resistant colonies.

15

#### C. In Vivo Properties Of $\Delta$ Ad.AAV1:

Viral DNA is labeled with BrdU during virus amplification to investigate cellular/nuclear vector uptake in situ. For transduction studies, confluent SKHep1 cells (a human endothelial cell line) are infected with 2000 genomes  $\Delta$ Ad.AAV1 or Ad.AAV1 per cell. BrdU tagged viral DNA is detected in 100% of nuclei at 3 hours post-infection for both viruses indicating efficient cellular and nuclear uptake of hybrid virus DNA.

#### RESULTS

25

The  $\Delta$ Ad.AAV1 vector transduces a cell in vitro forming G418 resistant colonies with an efficiency of 17 or 58%, after infection with an MOI of  $1 \times 10^3$  or  $1 \times 10^4$  genomes per cell, respectively. Approximately  $2 \times 10^4$   $\Delta$ Ad.AAV1 genomes are required to yield one stable transfectant. Since all stable colonies contain integrated  $\Delta$ Ad.AAV1 vector DNA, this number reflects the minimal integration frequency of  $\Delta$ Ad.AAV1 in SKHep1 cells which is comparable with that from rAAV (Rutledge, E. A. et al., 1997, *Journal of Virology*,

30

71:8429-36). The number of G418 resistant colonies does not necessarily represent the total frequency of integration events because not all integrated copies express neomycin phosphotransferase, due to chromosomal position effects or incomplete integration.

- 5 The absence of adenoviral gene products in  $\Delta$ Ad.AAV1 transduced cells at day 3 post-infection is demonstrated by immunofluorescence with antibodies to the major late proteins (hexon, fiber) and early proteins (DBP.E4-orf6). Expressed adenoviral proteins are detected only in cells infected with Ad.AAV1. The fact that cells infected with  $\Delta$ Ad.AAV1 do not express potentially cytotoxic adenoviral proteins is important.

10

- While an MOI of  $1 \times 10^4$  genomes per cell of the first generation vector Ad.AAV1 induce cytopathic effects in SKHep1 cells at day 3 p.i., no toxic side effects are observed when SKHep1 cells are infected with  $\Delta$ Ad.AAV1 at a dose of up to  $1 \times 10^8$  genomes per cell. Since the transduction efficiency is clearly dose dependent,  $\Delta$ Ad.AAV1 (which can be  
15 produced at titers of  $>5 \times 10^{12}$  genomes/ml) is able to stably transduce Ad5 permissive cell lines or tissues with a 100% efficiency without associated toxicity.

- Southern analysis indicates that  $\Delta$ Ad.AAV1 integrates randomly as head-to-tail tandem repeats into the host cell genome via the right AAV ITR, whereas the other junction with  
20 the chromosomal DNA is variable and occurs somewhere within the transgene cassette. In order to confirm the integrated status of  $\Delta$ Ad.AAV1 DNA, high-molecular-weight chromosomal DNA is separated by pulse field gel electrophoresis (PFGE), followed by Southern analysis with a SEAP specific probe (Figure 3). Undigested DNA from control SKHep1 cells give an endogenous SEAP signal that co-migrates with chromosomal DNA  
25 just below the well (lanes 1 and 5). No high-molecular weight episomal forms of  $\Delta$ Ad.AAV1 DNA are detected, whereas a distinct 35 kb band is visible in DNA from SKHep1 cells isolated 3 days after infection with first generation adenovirus, Ad.AAV1 (lanes 4 and 13). Digestion with EcoRI reveals the 4.4kb fragment, which is specific for integrated tandem copies of the AAV cassette (lanes 8 and 12). To eliminate the  
30 possibility that chromosomal DNA is trapped in the well, DNA samples are digested with intron-encoded endonucleases PI-SceI or I-CeuI (Gibco-BRL, Grand Island, NY) with a

sequence specificity or more than 11bp or 9bp respectively. Digestion with PI-SceI yields a >2mb endogenous SEAP signal in SKHep1 cells (lane 2) and an additional signal in the range of ~1mb in G418 resistant colonies transduced with  $\Delta$ Ad.AAV1 (lane 7). I-CeuI digestion results in a smear between 250-1000 kb in  $\Delta$ Ad.AAV1 transduced SKHep1-cells (lanes 10, 11) indicating random integration, whereas a high-molecular weight band specific for the endogenous SEAP gene is observed in control SKHep1 cells (lane 9).

One day after intraportal infusion of  $1 \times 10^{12}$   $\Delta$ Ad.AAV1 genomes in C57Bl/6 mice, BrdU labeled vector genomes can be detected in 85% hepatocytes (Lieber, A., et al. 1999. *J Virol* 73:9314-24). Hepatocellular DNA analysis performed at 2 months post-infusion reveals  $\Delta$ Ad.AAV1 DNA integrated with an average of 0.5 copies per cell into the mouse genome (Lieber, A., et al. 1999. *J Virol* 73:9314-24). To assess potential side effects of intraportal  $\Delta$ Ad.AAV1 infusion, serum glutamic pyruvic transaminase (SGPT), a sensitive marker for hepatocellular injury, is measured for 7 consecutive days post-infusion in combination with histological analysis of liver sections. No significant elevation in SGPT levels, or histological abnormalities are detected after intraportal infusion of  $1 \times 10^{12}$  or  $1 \times 10^{13}$   $\Delta$ Ad.AAV1 genomes, whereas infusion of the same dose of full-length Ad.AAV1 vector is associated with severe hepatotoxicity or fatal outcome. This suggests that the dose of  $\Delta$ Ad.AAV1 administered to mice can be increased to obtain higher transduction efficiencies in vivo without adverse side effects, which is not possible for first generation adenoviruses. Importantly,  $\Delta$ Ad.AAV1 transduced quiescent hepatocytes in vivo, which suggests that integration of hybrid vector DNA may not require cell proliferation. Recently, more detailed in vivo transduction studies with Ad.AAV1 and  $\Delta$ Ad.AAV1 have been performed in Balb/c mice to study whether the absence of adenoviral gene expression in cells infected with  $\Delta$ Ad.AAV1 can avoid an anti-viral immune response and can prolong vector persistence. In this mouse strain, vector DNA is cleared from the liver at 4-6 weeks after infusion with first generation adenoviruses, mostly due to a CTL response against viral proteins produced in transduced cells. Vector DNA is analyzed by genomic Southern Blot of hepatic DNA at 12 weeks

after infusion of  $1 \times 10^{12}$  genomes Ad.AAV1 or  $\Delta$ Ad.AAV1. At this time point, no vector specific signal is detectable in hepatic DNA from mice infused with the first generation vector Ad.AAV1, while  $\sim 0.3$  copies of  $\Delta$ Ad.AAV1 genomes per cell are present in livers of mice that received the hybrid vector, again indicating the superior in vivo properties of the hybrid vector.

#### D. Effects of Rep Coexpression on $\Delta$ Ad.AAV Integration

##### Rep expression after plasmid transfection:

In order to test whether Rep expression enhances site-specific integration of  $\Delta$ Ad.AAV1 in human cells, a series of Rep expression plasmids are constructed.

#### METHODS

The Rep ORF 68/78 (nt 285-2313) including the internal p19 and p40 promoters is obtained from pAAV/Ad (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) by digestion with BsaI/BsrI. This fragment deleted for the AAV p5 promoter is cloned via adapter linkers under RSV or PGK promoter in front of the bovine growth hormone polyadenylation signal (bPA) into pAd.RSV or pAd.PGK (Lieber, A., and Kay, M.A., 1996, *J. of Virology*, 70, 3153-3158; Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) correspondingly.

#### RESULTS

The resulting plasmids (pRSVrep, pPGKrep) are transfected into 293 cells or SKHep1 cells, most of the Rep proteins expressed from the heterologous promoters (RSV or PGK) are Rep 68 and Rep 78, while transfection of the rep gene under aP5 promoter (pAAV/Ad) results in predominant Rep 52/40 expression. Thus, transfection of pRSVrep and pPGKrep is more pronounced suggesting a strong transactivation of AAV promoters

by Ela which is produced in 293 cells. This result indicates that minimum expression of rep proteins is necessary to avoid interference with adenovirus replication.

Rep-mediated site-specific integration of  $\Delta$ Ad.AAV1.

5 The potential for site-specific integration is an important characteristic of the novel Ad.AAV vectors of the invention. In an embodiment of the invention, integration of the  $\Delta$ Ad.AAV is directed by co-infection with Ad AAV expressing the rep 78 protein to achieve site-specific integration in the AAVS1 site on human chromosome 19. For this type of site-specific integration to occur in cells other than 293 cells, E4 ORF6  
10 expression is required. The co-infection of  $\Delta$ Ad.AAV,  $\Delta$ Ad. rep 78, and  $\Delta$ Ad. E4-orf6 allows for site specific integration of the  $\Delta$ Ad.AAV transgene cassette. The  $\Delta$ Ad. rep78 and the  $\Delta$ Ad. E4-orf6 genomes are degraded soon after transduction, thus avoiding potential side effects. Site-specific integration is preferred over random integration,  
15 which is seen with rAAV and  $\Delta$ Ad.AAV, in order to reduce the risk of insertional mutagenesis.

A preliminary test can be performed to confirm the functional activity of Rep 68/78 expressed from pRSVrep to mediate site-specific integration of  $\Delta$ Ad.AAV1 (Figure 5 and  
20 6). Human SKHep1 cells are transfected with pRSVrep or control plasmid (pRSVbGal (Lieber, A., et al., 1995, *Human Gene Therapy*, 6, 5-11) (transfection efficiency was ~20%), followed by infection with  $\Delta$ Ad.AAV (2000 genomes per cell). Three days after infection, cells are trypsinized, embedded in agarose, lysed in situ, digested with I-CeuI (an intron-encoded endonuclease with a recognition sequence of more than 10nt),  
25 subjected to pulse field gel electrophoresis in 1% agarose gel, and analyzed by Southern Blot. Hybridization with a probe covering the AAVS1 integration site (1.7kb EcoRI/BamHI fragment from the chromosome 19 locus (Samulski, R. J. et al., 1991, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia) ) reveals an AAVS1-specific band (~240kb) in I-CeuI digested DNA from  
30 cells after control plasmid transfection (pCo) +  $\Delta$ Ad.AAV1 infection. An additional signal in the range of 280kb appears in rep expressing cells infected with  $\Delta$ Ad.AAV1

(pRSVrep + VAd.AAV1) indicating a site-specific insertion into the AAVS1 site in a certain percentage of cells. The presence of vector DNA in this 280kb band is confirmed by rehybridization of the same filter with a transgene (SEAP) specific probe. Randomly integrated  $\Delta$ Ad.AAV1 vector appears as a diffuse SEAP signal in the range 280-680kb  
5 (pCo+ $\Delta$ Ad.AAV1, pRSVrep+ $\Delta$ Ad.AAV1). The specific ~1.9 mb band on blots hybridized with the SEAP probe represents an I-CeuI fragment containing the endogenous human SEAP gene.

Incorporation Rep 68/78 function into hybrid vectors to stimulate site-specific integration  
10 Rep overexpression inhibits adenovirus DNA replication, prohibiting the generation of rep expressing Ad vectors using conventional strategies. To solve this problem, significant Rep 68/78 expression from the hybrid vector in virus producer (293) cells must be prevented while maintaining transient Rep expression in target cells (HSC) to mediate site-specific integration. Our hypothesis is that the specific structure of the  
15  $\Delta$ Ad.AAV hybrid virus can be used to bring the rep gene 68/78 into a transcriptionally active position under control of a HSC specific promoter only at late stages of virus replication in 293 cells. This will allow amplification of the hybrid vector in 293 cells, generating high titer virus which activates the incorporated Rep 68/78 functions only in HSC. The general outline of our strategy to produce Rep expressing hybrid vectors is  
20 illustrated in Figure 7. The rep/transgene cassette is assembled based on the left-hand shuttle plasmid used for recombinant adenovirus production. The gene encoding Rep 68/78 is cloned in 3' to 5' orientation in front of a transgene expression cassette flanked by AAV ITRs. Between the transgene cassette and the right AAV ITR an HSC-specific promoter is inserted with direction towards the adenoviral E2, E3, and E4 genes. The  
25 recombinant genome is produced by recombination in E.coli and transfection into 293 cells generates virus (Ad.AAV-rep). The specific structure of  $\Delta$ Ad.AAV with duplicated sequences flanking the AV ITRs is used to bring the rep gene into a transcriptionally active position under control of a HSC specific promoter only during late stages of viral DNA replication in 293 cells. During amplification of Ad.AAV-rep, the smaller genome  
30  $\Delta$ Ad.AAV-rep is formed and packaged into particles, which can be separated by ultracentrifugation in CsCl gradients. The specific structure of  $\Delta$ Ad.AAV-rep brings the

rep gene into 5'→3' orientation in relation to the HSC specific promoter, allowing rep transcription in target cells. After transduction of HSC with purified ΔAd.AAV-rep particles, rep expression is activated and mediates rescue of the AAV-ITR/transgene cassette from the adenoviral vector backbone and site-specific integration. The hypothesis is that Rep-mediated integration into AAVS1 occurs via the right or both AAV/ITRs causing the rep gene to become separated from the hepatocyte-specific promoter once the vector is integrated (Figure 7). Therefore, rep expression should be only transient without critical cytotoxic side effects on the host cell.

10 Promoters that can regulate rep expression:

Potential candidate promoters to drive rep expression with high specificity for HSC and minimal activity in 293 cells are the 454nt CD34 promoter (Krause, D.S., et al., 1997, *Experimental Hematology*, 25, 1051-1061; Yamaguchia, Y. et al., 1997, *Biochimica et Biophysica Acta.*, 1350:141-6), the 300nt HS 40 enhancer (Chen, H.L., et al., 1997, *Nucleic Acids Res.* 25, 2917-2922) or a 3kb CD34 enhancer (May, G. et al., 1995, *EMBO J.*, 14:564-74) in combination with an initiator, or the HIV LTR. An optimal promoter is selected based on studies of transient reporter gene expression after plasmid transfection in 293 cells and hepatocytes. All promoters to be tested are cloned in front of the human α<sub>1</sub>-antitrypsin (hAAT)- bovine growth hormone polyadenylation signal (bPA) into the adenoviral shuttle plasmid pCD2 (pAd.-hAAT). Promoter activity can be tested in transient plasmid transfection assays in CD34+ and 293 cells. The promoter with the highest hAAT levels in CD34+ or K562 cells and the lowest hAAT expression in 293 cells is selected for further studies. If high background expression in 293 cells from these promoters is seen, insulators to shield HSC-specific promoters from the Ela enhancer which is still present in Ad shuttle plasmids can be utilized.

Rep genes:

30 The large Rep 68/78 proteins are sufficient to mediate rescue and site-specific integration. Unregulated Rep 52 and Rep 40 expression from the AAV p19 promoter

located within the ORF of Rep 68 and 78 must be prevented because production of these smaller Rep proteins in 293 cells will affect cell viability and adenoviral DNA synthesis. To do this, constructs obtained from Surosky et al., containing a mutated Rep 52/40 start codon to express Rep 68 and 78 individually under CMV promoter can be used. The 293  
5 cells transiently expressing Rep68 or Rep 78 from these constructs can be coinfectd with  $\Delta$ Ad.AAV1 (infection 24 hours after pCMVRep transfection, MOI  $2 \times 10^5$  genomes/cell). Three days after  $\Delta$ Ad.AAV1 infection, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE as described earlier. Efficient Rep mediated excision of the AAV cassette and site-specific integration without flanking adenoviral  
10 sequences are expected and the plasmids pCMVRep68 or pCMVRep78 can be used as a source for the corresponding rep genes and clone them into hybrid vectors.

Vectors:

15 The rep/transgene cassette can be assembled based on pXCJL (Microbix, Toronto). A set of control hybrid vectors can be generated with the AAV-ITR-transgene cassette only without the rep gene. The recombinant Ad.AAV-rep genome can be generated by recombination of the left hand shuttle plasmids with pCD1, a pBHG10 (Microbix, Toronto) derivative, which contains the Ad5 genome deleted for the E1/E3 regions in  
20  $recA^+$  E. coli (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Compared to the standard technique based on plasmid recombination in 293 cells, this approach has the advantage that plaques with recombinant virus appear 3 times faster and the production of illegitimate recombinants is minimized. This allows efficient viral DNA amplification and packaging to occur before Rep expression reaches levels that are potentially  
25 inhibitory for adenoviral replication. The critical variables in maximizing the output of the vector deleted for all adenoviral viral genes are the initial multiplicity of infection and the time of harvesting. These parameters can be optimized for production of  $\Delta$ Ad.AAV-rep hybrid vectors. A number of  $\Delta$ Ad.AAV vectors can be constructed incorporating rep gene. Cryptic promoter and enhancer elements present in the 5'-342nt of the adenoviral  
30 genome can interfere with transgene expression from the heterologous promoters. This is crucial for the strategy to avoid rep expression from  $\Delta$ Ad.AAV-rep genomes in 293 cells.



To ensure efficient transgene expression, insulator fragments such as the chicken beta-globin insulator can be used with a selected promoter, constitutive or inducible.

Rep protein co-packaging:

5

As an alternative to producing hybrid vectors containing the rep 68/78 gene, studies are designed to see whether Rep protein can be co-packaged into  $\Delta$ Ad.AAV capsids and whether these co-packaged Rep molecules are sufficient to mediate rescue and site-specific integration of the AAV-ITR-transgene cassette. Our hypothesis is that the Rep  
10 68/78 binds to the Rep binding site (RBS) present in double-stranded  $\Delta$ Ad.AAV genome and that this complex is co-packaged into adenoviral capsids which are spacious enough to accommodate extra proteins. Based on protein/DNA ratio analysis performed previously in purified particles that only one 5.5kb  $\Delta$ Ad.AAV1 genome is packaged per capsid. This is confirmed by electron-microscopy of  $\Delta$ Ad.AAV1 particles, which reveals  
15 only spotted electron-dense staining associated with viral cores and extended free luminal space (see Figure 2).

293 cells are transfected with plasmids expressing Rep 68/78 under the CMV promoter and the kinetics of rep expression is determined by Western Blot with cell lysates  
20 collected at different time points after transfection. Next, these 293 cells are infected with Ad.AAV (MOI 1, 10, 100 pfu/cell) at specific time points after transfection of Rep plasmids depending on the Rep expression kinetics (e.g. 3, 6, 12, 24 . . . hours after transfection). It is important to time Ad.AAV infection exactly because viral DNA replication must be taking place or finished before Rep production reaches peak levels.  
25 In general, adenovirus DNA replication in 293 cells (infected with MOI 10) is maximal at 18 hours post-infection, followed by production of structural proteins, packaging of viral genomes, and breakdown of cellular membrane structures (which is concluded ~36-48h p.i.) (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; van der Vliet, B., 1995, *In* w. Doerfler, et al. (eds.) vol. 2  
30 p. 1-31, Springer-Verlag, Berlin). Viruses are collected 48h after infection and banded by CsCl ultracentrifugation. Viral material from purified bands corresponding to

ΔAd.AAV is lysed, DNase-treated (to liberate DNA associate Rep) and subjected to immunoprecipitation-Western Blot with Rep specific antibodies to detect co-packaged Rep. Based on theoretical calculations assuming that two Rep molecules bind per Ad genome, ~1-10ng Rep proteins is expected from Lysates of  $10^{10}$  particles, which is within the range of detectability by Western Blot. Alternatively, co-packaged Rep may be detected based on its functional activity to mediate rescue and site-specific integration of the AAVITR transgene cassette. To test whether functional Rep protein is co-packaged into hybrid vector particles, CsCl purified ΔAd.AAV1 particles generated in 293 cells co-expressing Rep after Ad/AAV1 infection (ΔAd.AAV1+Rep) can be used for transduction studies. Three days after ΔAd.AAV1+Rep infection of the human cell line K562, cellular DNA is analyzed for AAVS1-specific integration events by PCR and PFGE. If efficient Rep-mediated site-specific integration of excised AAV cassettes is successful, then other ΔAd.AAV+Rep hybrid vectors with β-Gal and SNori as transgenes can be produced.

Integration studies with Rep vectors in erythroid cells:

The hypotheses behind the rational of a rep-expressing hybrid vector (ΔAd.AAV-rep) are: (1) transient Rep co-expression from ΔAd.AAV-rep vectors can enhance site-specific vector integration in human cells and (2) integration occurs via the AAV ITR(s) without the rep gene, which is placed outside the AAV cassette, thus eliminating rep expression upon vector integration. To test hypothesis 1, transduction frequencies of ΔAd.AAV/rep versus ΔAd.AAV vectors can be compared based on the formation of G418 resistant colonies and quantify site-specific integration events at different time points after infection of human and mouse cells by PFGE and PCR. To test hypothesis 2, the structure of integrated vector in transduced cell populations and single clones can be delineated by Southern analysis and by sequencing of vector/chromosomal DNA junctions. These studies can be performed with ΔAd.AAV-rep, ΔAd.AAV, and ΔAd.AAV+Rep (copackaged protein) in human K562 or HEL (for AAVS1 integration) and mouse MEL cell lines.

Cells infected with  $\Delta$ Ad.AAV-SNori,  $\Delta$ Ad.AAV-SNori+Rep or  $\Delta$ Ad.AAV-Snori-rep can be subjected to G418 selection. The number of G418 resistant colonies determined after 4 weeks of selection in relation to the number of initially infected cells. The selection process for colonies that did not survive continued selection due to potential rep-mediated cytotoxicity or episomal vector expression can be monitored. If rep expression from  $\Delta$ Ad.AAV-SNori rep does not affect cell viability and proliferation, then more G418 resistant colonies should appear in  $\Delta$ Ad.AAV-SNori-rep and  $\Delta$ Ad.AAV-Snori+Rep. The structure of integrated vector can be determined by Southern Blot and sequencing of integration junctions.

To uncover a potential selection bias against rep producing cells after transduction with  $\Delta$ Ad.AAV/rep, site-specific and random vector integration events can be quantitated in cellular DNA isolated from cell populations at different time points after infection (e.g. 0.5, 1, 3, 7, 14 days). To do this, the techniques based on PFGE-Southern can be utilized.

It is expected that the signal(s) for AAVS1-specific integration in  $\Delta$ Ad.AAV/rep infected human cells increases during the first days after infection and then remains constant over time.

In a separate study, the integration status of vector DNA (analyzed by PFGE or PCR) and the number of integrated copies (analyzed by Southern Blot) with the expression level of  $\beta$ -galactosidase in single clones transduced with  $\beta$ -Gal hybrid vectors ( $\Delta$ Ad.AAV-BG,  $\Delta$ Ad.AAV-BG+Rep, or  $\Delta$ Ad.AAV-BG-rep) can be correlated. Together with data obtained in the studies described in the Specification, this allows assessment of whether transcriptional silencing is associated with site-specific vector integration into the AAVS1 site.

It is not clear a priori whether the specific Rep function for vector rescue, concatemerization, and integration can efficiently occur in non-S-phase or non-dividing cells. To test whether  $\Delta$ Ad.AAV<sup>fix</sup>,  $\Delta$ Ad.AAV or  $\Delta$ Ad.AAV-rep/+Rep vectors can integrate into non-dividing cells, transduction studies in cell cycle arrested cell cultures can be performed as described earlier.

## DISCUSSION

5 The establishment of stable cell lines expressing Rep 68/78 at detectable levels is not possible, which is probably due to rep mediated cytotoxicity. Therefore, it is not possible to perform long-term transduction studies (e.g. G418 selection or studies in single clones) in combination with ectopic rep expression. Moreover, due to the inhibitory effect of rep on adenovirus replication, it is currently not possible to generate adenoviral vectors expressing rep under the RSV or PGK promoter.

10

Taken together, this indicates that co-expressed Rep may stimulate site-specific transgene integration.

### E. A Detailed Study Of Transduction/Integration Of Hybrid Vectors In Erythroid Cell Lines:

15

In order to improve transduction and integration frequencies of the hybrid vectors into erythroid cell lines, a detailed study comparing various hybrid vectors have to be carried out as described below. The transduction studies are performed in K562 cells which is considered to be an adequate model to study gene transfer vehicles into erythroid cells (Floch, V., et al., 1997, *Blood Cells, Mol. and Diseases*.23, 69-87). The optimal vectors should be able to integrate into the cellular genome with a high frequency, determined by Pulse field gel electrophoresis (PFGE) and Southern blot as described in Example 4. In addition, the results from the following studies will serve to evaluate whether a given hybrid vector needs to be modified for site-specific integration in the host genome.

25

### Sequencing of integration junctions:

The ultimate proof for vector integration is the sequencing of junctions between SNori vector DNA and chromosomal DNA. Furthermore, this clarifies the question whether the AAV ITRs represent the substrate for integration. Specifically, DNA from clones with

30

known  $\Delta$ Ad.AAVSNori integration structure (analyzed by Southern Blot) digested with EcoRI, which does not cut within the SNori cassette. The resulting fragments are circularized and transformed into a specific *E. coli* strain (according to the protocol described by Rutledge and Russell (Rutledge, E. A. et al., 1997, *Journal of Virology*, 5 71:8429-36)). Kanamycin resistant bacterial clones should contain the integrated SNori cassette. Flanking chromosomal DNA in rescued plasmids can be sequenced with primers specific to the transgene.

To confirm vector integration in a small number of transduced cells, genomic DNA is 10 extracted and digested with EcoRI. EcoRI fragments are ligated to linkers containing a specific primer binding site and are then digested with NotI, religated and propagated in *E. coli*. Plasmid DNA from a representative number of bacterial clones is sequenced to determine the vector/chromosomal DNA junctions.

15 Dose dependent toxicity:

In order to test that the transduction frequency is dose-dependent and  $\Delta$ Ad.AAV vectors, which are devoid of all adenoviral genes, could be used to infect cells at higher doses with less cytotoxicity than first generation adenovirus, K562 cells are infected with 20 different MOIs ( $1-10^8$ ) of  $\Delta$ Ad.AAVBG and the first generation vector Ad.AAVBG (which contains the same  $\beta$ -Gal expression cassette). At day 4 post-infection, the total number of cells, the percentage of viable cells (based on trypan blue exclusion) and the percentage of X-Gal positive cells are counted. A fraction of infected cells are quantified for  $\beta$ -Gal expression using the Galacto-Light kit. The level of transgene expression is 25 expected to be comparable between the two vectors. K562 cells are predicted to tolerate higher doses of  $\Delta$ Ad.AAVBG better than Ad.AAVBG which express viral genes.

Integration frequency with and without G418 selection:

30 In order to investigate the integration frequency of the different vectors and to confirm that AAV ITRs present in double-stranded adenoviral DNA genomes can mediate vector

integration with a frequency comparable to rAAV vectors. integration studies are performed based on the formation of G418 resistant colonies with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.SNoriITR, and rAAVSNori after infection with  $2 \times 10^5$  and  $2 \times 10^6$  genomes per cell (Fig. 8). After infection, cells are plated in 96 well plates under limiting dilution and selected with G418 to estimate the frequency of formation of G418 resistant colonies. Another set of cells is plated without G418. A representative number of clones (w/ and w/o G418 selection) are expanded to  $>10^6$  cells (after 3-4 weeks of culture) and analyzed for the presence of viral DNA by Southern Blot as well as PFGE analysis to discriminate between episomal vector DNA and vector genomes stably associated with chromosomal DNA. This allows us to estimate the integration frequency of the different vectors, to assess the effect of G418 selection on integration, and to consider position effects on neo expression in calculating the total integration frequency. Integrated vector copies with a frequency of at least  $1 \times 10^{-4}$  is predicted only for  $\Delta$ Ad.AAVSNori and rAAVSNori. The total number of colonies may be lower in both the first generation vectors, Ad.AAVSNoriITR and Ad.SNori, due to the toxic effects of expressed adenoviral proteins; however, a higher integration frequency is predicted for the vector containing the AAV ITRs (Ad.AAVSNoriITR).

#### Kinetics of integration:

Compared to rAAV, the double-stranded nature of entering  $\Delta$ Ad.AAV genomes provides more protection against degradation. Furthermore, the synthesis of transcriptionally active double-stranded intermediates from single-stranded genomes, which is considered a limiting step in rAAV transduction, is not required in  $\Delta$ Ad.AAV transduction. Thus, the lag phase between infection and expression seen with rAAV vectors, which is causally linked to double-strand synthesis/integration may be shorter or absent in infections with  $\Delta$ Ad.AAV vectors. Furthermore, it was demonstrated earlier that a 9kb mini-adenoviral genome packaged into adenoviral particles is only short lived and completely degraded by day 3 post-infusion. In contrast, transduction with the 5.5kb  $\Delta$ Ad.AAV1 (Figure 1) genome allows for long-term expression, suggesting that either

AAV ITRs can stabilize the viral genome as an episome until it is integrated or integration occurs shortly after infection.

The status of vector DNA can be examined in K562 cells at different time points after infection with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.AAVSNoriITR, or rAAVSNori (MOI  $2 \times 10^5$ ). Infected cells are harvested at 1 hour, 5 hours, 1 day, 3, 7, and 14 days after infection and chromosomal DNA is analyzed by PFGE followed by hybridization with a transgene specific probe. This technique allows us to distinguish between episomal vector DNA, which appears as a distinct 5.0kb band and integrated DNA. Furthermore, extra chromosomal high-molecular weight vector concatemers can be detected. In the case of random integration, after digestion of chromosomal DNA with I-CeuI or PI-SceI, vector-specific signals in the range of 1-2mb should be seen. The intensity of episomal and integrated vector signal is quantified for each time point using phosphorimager analysis. This gives information about the kinetics of hybrid vector integration in a population of infected K562 cells and the intracellular stability of hybrid vector genomes.

Structure of integrated vector DNA and integration junctions with chromosomal DNA:

$\Delta$ Ad.AAV1 integrates as concatemer/s randomly into host DNA as shown previously. How many vector copies are present in one concatemer and whether the extent and the kinetics of tandem-formation are dose dependent still remain unclear. Another unanswered question is how  $\Delta$ Ad.AAV integrates: whether one or both ITRs are involved, whether the integrated ITRs are still intact, and whether adenoviral sequences integrate as well. These issues are important for the strategy to include rep genes into the hybrid vector genome. Moreover, if intact AAV ITRs are present within integrated vector copies, helper virus (adenovirus or HSV) infection in vivo may mobilize the integrated AAV-ITR vector cassette and affect stability of transgene expression.

To answer these questions, K562 cells can be infected with  $\Delta$ Ad.AAVSNori, AdSNori, Ad.AAVSNoriITR, and rAAVSNori at MOIs  $2 \times 10^5$ ,  $2 \times 10^6$ , or  $2 \times 10^7$  genomes per cell. Infected cells are plated in 96 well plates in the presence or the absence of G418. The

latter is included because G418 may cause amplification of integrated vector DNA (Rutledge, E. A. et al., 1997, *Journal of Virology*, 71:8429-36). Genomic DNA from isolated clones can be analyzed by regular Southern Blot as described in the Examples Section to confirm the presence of vector concatemers and calculate the number of integrated vector copies. More informative is the sequencing of integrated vector copies and their junctions with chromosomal DNA. The structure of integration junctions can be delineated using the role of AAV ITRs in vector integration and the extent of insertional mutagenesis after transduction. This data provides information about the potential risks of hybrid vector used in clinical trials.

Transduction of cell cycle arrested cells:

The ultimate target for the hybrid vectors described in the Specification are quiescent hematopoietic stem cells. We hypothesize that the double-stranded nature of  $\Delta$ Ad.AAV genomes and specific nuclear import mechanisms may allow for the transduction of non-dividing cells. This is in part supported by the transduction studies with  $\Delta$ Ad.AAV1 in quiescent hepatocytes in vivo. To confirm this data, primary fibroblasts can be forced to enter the G<sub>0</sub> phase by serum/growth factor starvation before infection with the hybrid vectors according to a protocol described by Russell (Russell, D. et al., 1995, *PNAS*, 92:5719-23). Cells are maintained for three days after infection under serum/growth factor deprivation. At this time point, genomic DNA is isolated and analyzed for integration events by PFGE in comparison with growing cells. Another series of integration studies can be carried out on K562 cells arrested in the G<sub>1</sub>/S phase of the cell cycle with aphidicolin (added 1 day before and maintained several days after infection with hybrid vectors - depending on the integration kinetics studies described earlier). To investigate whether DNA damaging agents increase the transduction frequency of hybrid vectors, cell-cycle-arrested K562 cells or primary fibroblasts can be treated with cisplatin or <sup>3</sup>H-thymidine prior to virus infection according to a protocol described by Alexander and Russell (Alexander, I. E. et al., 1994, *J. Virol.*, 68:8282-87; Russell, D. et al., 1995, *PNAS*, 92:5719-23). Furthermore, the effect of chromosomal DNA decondensation on the transduction efficiency of hybrid vectors can be studied in arrested



cells after treatment with puromycin, staurosporin, Hoechst 3328, distramycin, or vandate.

#### **F. Improvements in $\Delta$ Ad.AAV Production and Purification**

5 To inhibit packaging of full-length genomes a modified form of I-Sce I, a yeast mitochondrial intron-endonuclease with a non-palindromic 18-bp recognition sequence is expressed in 293 cells. Constitutive expression of this enzyme in mammalian cells is not toxic, possibly due to either the lack of I-SceI sites in the genome or sufficient repair of  
10 them (Rouet P. et al, 1994, *PNAS*, 91:6064-8). The yeast I-Sce I is modified with an SV40 T-antigen nuclear localization signal and an optimal Kozak sequence to enhance its functionality in mammalian cells (Rouet P. et al, 1994, *PNAS*, 91:6064-8). For another yeast endonuclease it was shown that a recognition site within an transduced Ad genome was efficiently (30% of all transduced genomes) when expressed in human A549 cells.  
15 Importantly, the expression of E4 ORF6 and ORF3 expressed from the transduced Ad genome inhibited double-strand break repair mediated by the endonuclease (Nicolas, A. L. et al, 2000, *Virology*, 266:211-24). This is consistent with the observations by others where these E4 proteins prevent concatemerization of the viral genome (Boyer, J. et al, 1999, *Virology*, 263:307-12). Based on this, packaging of full-length virus containing a  
20 I-SceI recognition site is reduced in 293 cells constitutively expressing I-Sce I. The 18mer I-Sce site is inserted into the E3 region of the Ad.IR vectors. These vectors are generated and amplified in 293 cells followed by a large-scale infection of 293 cells expressing I-SceI. Alternatively, an expression cassette for the endonuclease XhoI is inserted into the E3 region of Ad.IR or Ad.AAV vectors. The XhoI gene will be  
25 modified for optimal function in mammalian cells. Vectors expressing XhoI are generated and amplified in 293 cells expressing the XhoI isoschizomer PaeR 7 methyltransferase (PMT) (Nelson, J. E. et al, 1997, *J. Virol.*, 71:8902-7), which mediates the addition of a methyl group onto the N6 position of the adenine base of XhoI sites, CTCGAG. This protects the viral and cellular genome from XhoI cleavage. Methylated  
30 Ad vectors are produced at high titers.  $\Delta$ Ad.AAV vectors are then obtained by large-scale infection of 293 cells with the Ad.AAV-XhoI vectors. At this stage the viral

genome is not methylated and is digested at the XhoI sites. XhoI sites present within the transgene cassette are deleted by site-directed mutagenesis without altering the amino acids sequence. (XhoI is accumulated only at late stages in virus replication and should act only upon a large part of Ad DNA when replication is completed. In addition, ultracentrifugation optimizes the separation between  $\Delta$ Ad.IR and  $\Delta$ Ad.IR particles (Blague, C. et al., 2000, *Blood*, 95:820-8).

## EXAMPLE II

### MODIFIED FIBER PROTEIN

#### A. Test the Infectivity of Different Human or Animal Serotype on Human Bone Marrow Cells.

Since the amino acid sequence of the fiber knob region varies considerably among the ~50 known serotypes, it is thought that different adenovirus serotypes bind to different cellular receptor proteins or use different entry mechanisms (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14; Defer, M., et al., 1993, *J. of Virology*, 64, 3661-3673). Although most adenoviruses contain RGD motifs in the penton base proteins, there are a number of serotypes (e.g. Ad 40, 41) without this conserved sequence. These types may use integrin  $\alpha$ v-independent pathways for virus internalization (Davison, A.J., et al., 1993, *J. Mol. Biol.*, 234, 1308-1316; Mathias, P. et al., 1994, *Journal of Virology*, 68:6811-14). To test whether other Ad serotypes can infect stem cell subpopulation present in human bone marrow, studies with a series of different human Ad serotypes and animal viruses can be performed (see Table II). As a means to verify efficient transduction with Ad serotypes, viral DNA is tagged before infection and the presence of viral genomes in the nuclei of transduced cells is investigated. Furthermore, whether viral DNA is replicated in transduced cells can be analyzed as indirect proof for early viral gene expression. A direct detection of expressed viral proteins is impossible due to the unavailability of antibodies against all the serotypes included in this study. Simultaneously with the infection assay, transduced

human bone marrow cells can be analyzed for morphological and immunohistochemical features characteristic of HSC or progenitor subpopulations. For retargeting, serotypes which are able to infect CD34+ subsets of bone marrow cells at the lowest MOI are selected. As the next step, the fiber gene is PCR-cloned from serotypes with potential  
5 HSC/CD34+ tropism and inserted into standard shuttle plasmids for Ad5 vector generation replacing the Ad5 fiber gene using an E. coli recombination system (Figure 9).

## METHODS

10

### Cells and viruses:

HeLa (human cervix carcinoma, ATCC CCL-2.2), CHO (chinese hamster ovary, ATCC CCL-61), K562 (human hematopoietic, ATCC 45506), HEP-2 (human larynx carcinoma,  
15 ATCC CCL-23), 293 (human embryonic kidney, Microbix, Toronto Canada) cells were maintained in DMEM, 10% FCS, 2 mM glutamine, and Pen/Strep. Culture media for CHO cells was supplemented with 200µM asparagine and 200µM proline. Human CD34+-enriched bone marrow cells were purified from peripheral blood after mobilization using MiniMACS VS<sup>+</sup> separation columns (Miltenyi Biotec, Auburn, CA),  
20 according to the manufacturer's instructions. Aliquots were stored in liquid nitrogen. Sixteen hours before the experiment, cells were recovered from the frozen stock and incubated overnight in IMDM media, supplemented with 20% FCS, 10<sup>-4</sup> M β-mercaptoethanol, 100 µg/ml DNaseI, 2 mM glutamine, 10 U/ml IL-3, and 50 ng/ml stem cell factor (SCF) or 2 ng/ml thrombopoietin (Tpo). The purity of CD34+ preparations  
25 was verified by flow cytometry and was consistently greater than 90%.

### Flow cytometry:

Adherent cells (CHO, HeLa) grown in non-tissue culture treated 10 cm dishes (Falcon,  
30 Franklin Lakes, NJ) were detached by treatment with 1mM EDTA and washed three times with wash buffer (WB), consisting of PBS supplemented with 1% FCS. Cells

grown in suspension (K562, CD34+) were washed three times with WB. After washing, cells were resuspended in WB at  $2 \times 10^6$  cells/ml.  $2 \times 10^5$  cells were incubated in WB for 1 h at 37°C with monoclonal antibodies specific for  $\alpha_v$ -integrins [L230, ATCC: HB-8448, (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795) (1/30 final dilution),  
5 CAR [RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652) (1/400 final dilution)], or BrdU [(Amersham, Arlington Heights, IL) (1/100 final dilution)]. Subsequently, cells were washed with WB, and incubated with fluorescein isothiocyanate (FITC)-labeled horse anti-mouse IgG antibodies [(Vector Labs., Burlingame, CA) (1/100 final dilution)] or phycoerythrin (PE)-  
10 labeled goat anti-mouse IgG antibodies [(Calbiochem, La Jolla, CA) 1:100 dilution] for 30 min at 4°C. After incubation with secondary antibodies, cells were washed two times with WB and  $10^4$  cells per sample were analyzed in duplicate by flow cytometry.

For the analysis of CD34 and c-kit expression on transduced CD34+-cells and for  
15 fluorescent activated cell sorting (FACS), purified human CD34+ cells were incubated with phycoerythrin(PE)-conjugated anti-CD34 monoclonal antibodies (Becton-Dickinson Immunocytometry Systems, San Jose, CA) or with PE-labeled anti-CD117 (c-kit) monoclonal antibodies (MAb 95C3, Immunotech, Beckman Coulter, Marseille, France) according to the manufacturer's protocol followed by flow cytometry analysis. All  
20 analyses and sortings were performed on a FACStar Plus flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with 488nm argon and 633 nm HeNe lasers. For analysis of c-kit expression and FACS purification of CD34+/c-kit+ cells, SCF was not added to the media during culturing of CD34+ cells.

## 25 RESULTS

### CAR/ $\alpha_v$ -integrin expression on test cells:

It is generally accepted that CD34+ cells possess bone marrow repopulating activity.  
30 Therefore, we used human CD34+ cells as the target for our studies towards identifying Ad serotypes with HSC tropism and constructing new viral vectors. Studies were

performed on mobilized, CD34-positive, peripheral blood cells from one donor under conditions which are known to retain CD34+ cells in a quiescent stage (Leitner, A., et al. 1996. *Br.J.Haematol.* 92:255-262; Roberts, A. W., Metcalf, D. 1995. *Blood.* 86:1600-1605). More than 90% of purified cells were CD34 positive by flow cytometry.

5 Furthermore, we included into our Ad tropism studies the cell line K562, which is considered to be an adequate model for studying gene transfer into human hematopoietic cells (McGuckin, et al. 1996. *British Journal of Haematology.* 95:457-460). HeLa cells, which are readily infectible by Ad5, and CHO cells, which are refractory to Ad5 infection (Antoniu, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9), were used as positive and

10 negative control cell lines, respectively.

For Ad5, both, binding to the primary receptor and to  $\alpha_3\beta_5$  and  $\alpha_v\beta_5$  integrins are important for high efficiency infection of target cells. The expression of CAR and  $\alpha_v$  integrins on test cells was analyzed by flow cytometry using monoclonal antibodies

15 against CAR (RmcB (Bergelson, J. M., et al. 1997. *Science.* 275:1320-1323; Hsu, K.-H., L., et al. 1988. *J. Virology.* 62:1647-1652)) and  $\alpha_v$  integrins (L230 (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621)) (Figure 10). As expected, nearly all HeLa cells expressed high levels of CAR and  $\alpha_v$ -integrins, whereas CHO cells lacked significant CAR and  $\alpha_v$ -integrin expression. Fifteen and 77% of K562 cells expressed CAR and  $\alpha_v$ -

20 integrins, respectively. Only ~6% of the CD34+ cells used in our studies expressed CAR and 17% were positive for  $\alpha_v$ -integrins. Notably, the preparation of CD34+ cells represents a mixture of different cell types. The absent or low expression of primary and secondary Ad5 receptors on non-cycling human CD34+ cells is in agreement with previous reports (Huang, S., et al. 1996. *J. Virology.* 70:4502-4508; Neering, S. J., et al.

25 1996.. *Blood.* 88:1147-1155; Tomko, R. P., et al., 1997. *Proc. Natl. Acad. Sci. USA.* 94:3352-3356).

Infection assay using wild-type Ad5 and K562 cells:

30 The presence of viral DNA in the nucleus of infected cells is an indirect means to demonstrate efficient virus binding, internalization, and nuclear import. Nuclear

localization of the viral genome is a prerequisite for transgene transcription and integration. Two techniques are utilized to tag viral DNA for in situ analysis. To optimize the infection assay, wild-type Ad5 virus and K562 cells which are permissive for Ad5 infection can be used. The first protocol (Challberg, S.S. and Ketner, S. 1981, *Virology* 114, 196-209), is based on  $^{32}\text{P}$ -labeling of viral DNA. During amplification of wild-type Ad5 and A549 cells,  $^{32}\text{P}$ -phosphate (40 $\mu\text{Ci/ml}$ ) is added to phosphate-free medium. After development of CPE,  $^{32}\text{P}$ -tagged virus is harvested, banded in CsCl gradients, and titered on HeLa cells according to standard protocols. To simulate the conditions for infection of human bone marrow cells, K562 cells are incubated in suspension with a MOI of 1, 10, or 100 of  $^{32}\text{P}$ -Ad5 for 2, 4, 6, or 8 hours under agitation at 37°C. This covers the time period necessary for adsorption, internalization, and nuclear import. After washing, cells are fixed either transferred to microscopy slides using cytospin or embedded in paraffin and sectioned (according to protocols from VECTOR labs, Burlingham, CA). The latter has the potential advantage that multiple consecutive sections (5 $\mu\text{m}$ ) of the same cell can be analyzed by different methods (e.g. for  $^{32}\text{P}$  tagged viral DNA, for specific histological staining, for immunofluorescence), which allows for correlating infection with a particular cell type present in the bone marrow. Cells are incubated in a Kodak NTB-2 photo emulsion for autoradiography. The exposure time can be optimized to minimize background or non-nuclear localized signals. A dose and time dependent appearance of nuclear silver grains is expected under the optimized conditions. Since  $^{32}\text{P}$ -phosphate can label viral proteins as well, a cytoplasmic background signal might appear. To facilitate detection, immunofluorescence with HSC specific antibodies on sections can be performed. As an alternative method, a BrdU-labeling technique for viral DNA can be used (Lieber, A., et al. 1999. *J Virol* 73:9314-24; Lieber, A. et al., 1996, *Journal of Virology*, 70:8944-60). In this case, different amounts of BrdU are added to the A549 culture medium during wtAd5 virus propagation. BrdU labeled viral DNA can be detected with monoclonal antibodies specific to BrdU. The signal can be enhanced using layers of species-specific polyclonal antibodies in combination with biotin/avidin and a fluorescent marker. BrdU tagged viral DNA can be detected on cytopspins of bone marrow cells together with cell surface markers by double or triple immunofluorescence.

## DISCUSSION

5 The interaction of selected Ad serotypes with CD34+ cells was tested. As a result of this screening we constructed a first-generation, Ad5-based vector whose fiber was substituted with the fiber derived from Ad35. We demonstrated that this capsid modification allowed for efficient viral transduction of potential HSCs by the corresponding chimeric Ad vectors.

10 All tropism and transduction studies were performed with non-cycling CD34+ cells, which are thought to include HSCs. The quiescent stage of CD34+ cells purified from mobilized blood is important because induction of cell proliferation is associated with a loss of the ability to reconstitute hematopoiesis and with changes in the spectrum of cellular receptors (Becker, P. S., et al. 1999. *Exp. Hematol.* 27:533-541). It is known that  
15 treatment of hematopoietic cells with cytokines or growth factors changes the expression of specific integrins including  $\alpha_v$ -integrins, which would ultimately alter the susceptibility of cells to Ad infection or may effect viability of infected cells (Gonzalez, R., et al. 1999. *Gene Therapy*. 6:314-320; Huang, S., et al. 1995. *J. Virology*. 69:2257-2263). Another fact that complicates the interpretation of transduction studies is the  
20 extraordinary heterogeneity of CD34+ cells in regards to morphology and function.

### B. Screening different adenoviruses to establish tropism to HSC.

The ATCC provides more than 70 different human or animal adenoviruses (see Appendix  
25 I). A collection of 15 human serotypes and 6 animal adenoviruses (see Table II) are selected based on the following criteria: (i) availability of the complete genome sequence or fiber sequence from the NIH gene bank (ii) CAR receptor usage absent or unknown, (iii) different subgroups, and (iv) moderate or low tumorigenicity (Shenk, T., 1996, *In* B. N. Fields, et al. (eds.), *Fields Virology*, vol. 2 Lippincott-Raven Publisher, Philadelphia).  
30 However, any serotype shown in the Appendix hereto can be used for the invention described. Animal viruses are included in the infectivity assay because this may provide a

means to circumvent the pre-existing humoral immunity against human Ad5 fiber, which represents a critical obstacle for clinical trials with Ad vectors.

## METHODS

5

### Viruses:

The following human adenovirus serotypes were purchased from the ATCC: 3 (VR-3), 4 (VR1081), 5 (VR-5), 9 (VR1086), 35 (VR-716) and 41 (VR-930). Adenovirus No. VR-  
10 716 was purchased from ATCC labeled as serotype 34, however it was found to be serotype 35 upon sequencing of the fiber region. For amplification, the corresponding Ads were infected onto HeLa, 293, or HEP-2 cells under conditions that prevented cross-contamination. Virus was banded in CsCl gradients, dialyzed and stored in aliquots as described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).  
15 Plaque titering was performed as follows: Confluent 293 cells plated in 6-well plates were incubated for 24 hours with virus in a total volume of 1ml. Two weeks after infection, plaques were counted on cultures overlayed with 1% agarose/MEM/10% FCS.0

### 20 EM studies:

CsCl-banded Ad stocks were thawed and diluted with 0.5% glutaraldehyde. Grids were prepared as described earlier (Mittereder, N., et al. 1996. *J. Virology*. 70:7498-7509). After staining with 2% methylamine tungstate (Nanoprobe, Stony Brook, NY), the  
25 carbon-coated grids were evaluated and photomicrographed with a Phillips 410 electron microscope, operated at 80 kV (final magnification 85,000x). For each particular Ad serotype, the number of morphologically deficient viral particles per 100 was counted in five random fields.

## 30 RESULTS



Electron microscopy:

Little is known about the stability of particles from serotypes other than Ad5. Since the intactness of viral particles was crucial for comparative interaction studies, virions from the serotypes specified above were analyzed by electron microscopy (EM). EM studies of negative contrast stained Ad suspensions demonstrated that the percentage of defective particles (loss of icosahedral shape or luminal staining) did not exceed 5% indicating that serotype preparations had comparable qualities. Representative EM photographs are shown for Ads 5, 9, and 35 (Figure 11).

Serotype screening:

It is thought that different Ad serotypes bind to different cellular receptor proteins and use different entry mechanisms (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Mathias, P., et al. 1994. *Journal of Virology*. 68:6811-6814). A set of human adenoviruses was obtained from the ATCC to be tested for tropism to CD34+ cells. These included serotypes 3, 4, 5, 9, 35, and 41 representing different subtypes (Table 1). We believed that these serotypes would use different cellular attachment and internalization strategies due to differing lengths of fiber shafts (Chroboczek, J., et al. 1995. *Adenovirus fiber*, p. 163-200. In a. P. B. W. Doerfler (ed.), *The molecular repertoire of adenoviruses*, vol. 1. Springer Verlag, Berlin; Roelvink, P. W., et al. 1998. *J. Virology*. 72:7909-7915), the presence or absence of RGD motifs within the penton base, and differing tissue tropism. The relatively little characterized Ad35 was selected because it was found in immunocompromised hosts, particularly in bone marrow recipients (Flomenberg, P., et al. 1994. *Journal of Infectious Diseases*. 169:775-781; Flomenberg, P. R., et al. 1987. *Journal of Infectious Diseases*. 155:1127-1134; Shields, A. F., et al. 1985 *New England Journal of Medicine*. 312:529-533). The latter observations prompted us to believe that bone marrow cells are among the natural reservoirs for Ad35.

TABLE II

Human and animal adenoviruses with potential interest for the invention

| Adeno virus | Human/ Group B              | Human/ Group D     | Human/ Group F | Avian     | Bovine | Canine | Ovine | Swine | Mouse |
|-------------|-----------------------------|--------------------|----------------|-----------|--------|--------|-------|-------|-------|
| Sero type   | <u>3,7</u> ,11,16, 21,34,35 | 8,15,17, 19,28,3 7 | <u>40, 41</u>  | CELO, EDS | 3      | 1,2    | 5     | 4     | 1     |

The underlined serotypes use CAR independent pathways for cell entry.

- 5 For amplification, the corresponding adenovirus stocks can be infected onto HeLa or A549 cells such that at a given time only one virus type is handled in a separate laminar flow hood and cultured in Hepa-filtered bottles, preferentially in separate CO<sub>2</sub> incubators to avoid cross-contamination. During propagation, viral DNA is tagged using one of the techniques described earlier. Viral DNA can be isolated from purified particles. The
- 10 XhoI restriction pattern is analyzed for methylated and unmethylated viral DNA by Southern blot using the full genome of the corresponding virus type as a radioactive probe.

## DISCUSSION

15

- Although it was reported earlier by slot-blot assay that fiber knobs derived from 2, 9, 4, and 41L can bind to CAR (Roelvink, P. W., et al. 1998. *J. Virology*. 72:7909-7915), it is not clear whether this binding occurs with an affinity that is physiologically relevant and whether this would confer cell entry. Furthermore, as shown for the Ad5 interaction
- 20 between the penton and integrins, a secondary receptor is required to induce virus internalization. We demonstrated that different serotypes interacted differently with the K562 or CD34+ target cells. Ad5, Ad4, and Ad41 were not able to efficiently attach to and be internalized by K562 and CD34+ cells. Although Ad4 belongs to a separate subgroup (E), it is thought that Ad4 represents a natural hybrid between subgroup B and
- 25 C viruses with a fiber related to Ad5 (Gruber, W. C., et al. 1993. *Virology*. 196:603-611).

Therefore, it was not surprising that Ad4 has binding properties similar to Ad5. The subgroup F serotype Ad41 has been shown to contain distinct fibers, a long shafted and a short-shafted fiber allowing for different cell entry pathways (Tiemessen, C. T., Kidd, A.H. 1995. *J. Gen. Virol.* 76:481-497). The Ad41 penton base does not contain RGD motifs suggesting that this virus may use  $\alpha_v$ -integrin independent pathways for cell entry. However, these features did not improve interaction with CD34+ cells. Ad9, Ad3, and Ad35 did interact with CD34+ cells more efficiently than Ad5. Out of all the serotypes tested, Ad35 demonstrated the most efficient attachment and internalization with K562 and CD34+ cells. Although the short-shafted Ad9 can bind to CAR, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Therefore, the low level of  $\alpha_v$ -integrin expression on certain subset/s of CD34+ cells may account for the observed susceptibility to Ad9.

### C. Attachment and Internalization of the Ad serotypes to K562 and CD34+ cells.

#### METHODS

##### Labeling of Ads with [<sup>3</sup>H]-methyl thymidine:

Serotypes were labeled with [<sup>3</sup>H]-methyl thymidine as described in detail elsewhere (Roelvink, P. W., et al. 1996. *J. Virology.* 70:7614-7621). Briefly, 5x10<sup>7</sup> HeLa or 293 cells were grown in 175 sq. cm flasks with 15 ml DMEM/10% FCS and infected with wild type adenovirus at a MOI of 50 or higher. Twelve hours post-infection, 1 mCi of [<sup>3</sup>H]-methyl thymidine (Amersham, Arlington Heights, IL) was added to the media and cells were further incubated at 37°C until complete CPE was observed. Then, cells were harvested, pelleted, washed once with cold PBS, and resuspended in 5 ml PBS. Virus was released from the cells by four freeze-thaw cycles. Cell debris was removed by centrifugation and viral material was subjected to ultracentrifugation in CsCl gradients and subsequent dialysis as previously described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology.* 70:8944-8960). Virus purification and dialysis removed unincorporated radioactivity. Wild type Ad particle concentrations were determined

spectrophotometrically by measuring the OD<sub>260</sub>, utilizing the extinction coefficient for wild-type Ad5  $\epsilon_{260} = 9.09 \times 10^{13}$  OD ml cm virion<sup>-1</sup> (Maizel, J. V., et al. 1968. *Virology*. 36:115-125). The virion specific radioactivity was measured by a liquid scintillation counter and was always in the range of  $1 \times 10^{-5}$  to  $1 \times 10^{-4}$  cpm per virion. For selected  
5 variants, the fiber gene was PCR amplified and sequenced to ensure identity and the absence of cross-contamination.

Viral DNA tagged with methylase and test for replication by genomic Southern blots:

10 To ultimately confirm transduction, a protocol to detect adenoviral replication in infected cells can be established. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. A site-specific methylation strategy is utilized to monitor viral DNA replication within infected cells (Nelson, J. et al., 1997, *Journal of Virology*, 71:8902-07). Methylation marked adenovirus can be produced by the addition of a  
15 methyl group onto the N6 position of the adenine base of XhoI sites, CTCGAG, by propagation of the virus in HeLa or A549 cells expressing the XhoI isoschizomer PaeR7 methyltransferase (PMT) (Kwoh, T.J., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 7713-7717). It is known that methylation does not affect vector production but does prevent cleavage by XhoI. Loss of methylation through viral replication restores XhoI cleavage  
20 and can be detected by Southern blots of genomic DNA from infected cells in comparison to native, non-methylated, viral genomes.

Attachment and internalization assays:

25 These studies were performed based on a protocol published elsewhere (Wickham, T. J., et al. 1993. *Cell*. 73:309-319). In preliminary experiments, we found that labeled Ad5 virions reached equilibrium in attachment to HeLa cells after 45 min at 4°C with an MOI of 400 pfu per cell. For attachment studies,  $3.5 \times 10^5$  cells were incubated for one hour on ice with equal amounts of [<sup>3</sup>H]-labeled adenovirus OD particles equivalent to an MOI of  
30 400 pfu/cell for Ad5 in 100 µl of ice-cold adhesion buffer (Dulbecco's modified Eagle's medium supplemented with 2 mM MgCl<sub>2</sub>, 1% BSA, and 20 mM HEPES). Next, the cells

were pelleted by centrifugation for 4 min at 1000 x g and washed two times with 0.5 ml ice-cold PBS. After the last wash, the cells were pelleted at 1500 x g, the supernatant was removed, and the cell-associated radioactivity was determined by a scintillation counter. The number of viral particles bound per cell was calculated using the virion specific radioactivity and the number of cells. To determine the fraction of internalized [<sup>3</sup>H]-labeled adenoviral particles, cells were incubated on ice for one hour with the corresponding virus, washed with PBS as described above, resuspended in 100 µl adhesion buffer, and then incubated at 37°C for 30 min. Following this incubation, cells were diluted 3-fold with cold 0.05% trypsin-0.5mM EDTA solution and incubated at 37°C for an additional 5-10 min. This treatment removed 99% of attached radioactivity. Finally, the cells were pelleted at 1500 x g for 5 min, the supernatant was removed, and the protease-resistant counts per minute were measured. This protocol minimizes the possibility that the internalization data were affected by receptor recycling (Rodriguez, E., Everitt, E. 1999. *Arch. Virol.* 144:787-795). Nonspecific binding of Ad particles to cells on ice was determined in the presence of 100-fold excess of unlabeled virus. This value routinely represented less than 0.1 % of viral load.

## RESULTS

### Attachment of Ad particles to target cells and internalization:

The selected serotypes were metabolically labeled with [<sup>3</sup>H]-thymidine, which is incorporated into viral DNA during replication. Adsorption and internalization can be experimentally dissociated by taking advantage of the observation that at low temperature (0-4°C) only virus cell attachment occurs, whereas internalization requires incubation at higher temperatures. The number of particles adsorbed or internalized per cell was calculated using the virion-specific radioactivity and used to quantify interaction of Ads 3, 4, 5, 9, 35, and 41 with CD34+, K562, HeLa and CHO cells (Figure 12). The serotypes varied significantly in their ability to attach to and to be internalized by the different cell lines. For Ad5, the degree of attachment to the cell lines tested correlated with the level of CAR expression. In CHO cells, which were previously shown to be refractory to Ad5

infection, the level of attachment and internalization was about 50-70 viral particles per cell. This number was hereafter assumed negative in terms of susceptibility of a given cell type for Ad5. Interaction of the other serotypes with CHO cells was not significantly higher indicating that corresponding receptor/s were absent on CHO cells. All serotypes tested interacted with HeLa cells; with Ad3 and Ad35 being the most efficient variants. The presence of distinct Ad3 and Ad5 receptors on HeLa cells was demonstrated previously (Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857). Ads 4, 5, and 41 did not bind to K562 cells. In contrast, Ad9 as well as the members of subgroup B, Ad3 and Ad35, efficiently interacted with K562 cells with Ad35 having the highest number of adsorbed and internalized particles. Compared to Ad5, about 25 times more Ad35 particles were attached and three-fourth of these were internalized by K562 cells. Viral interactions with CD34+ cells were generally weaker. Among the serotypes tested, only Ad9 and Ad35 were significantly internalized by non-cycling CD34+ cells. Internalization of Ad9 and Ad35 was, respectively, four and eight times more efficient than for Ad5 particles. The number of Ad35 virions internalized by CD34+ cells was almost half of that seen for Ad5 in HeLa cells, which can be readily infected with Ad5 based vectors.

Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into HeLa, 293, and CHO cells:

HeLa and 293 cells expressing high level of primary and secondary receptors for human adenoviruses are used as a positive control for virus attachment and internalization. As a negative control CHO cells are used. CHO cells do not express the primary adenoviral receptor at a detectable level, and are therefore refractory for adenoviral infection. For attachment studies, these adherent cell lines are detached from 10 cm dishes with PBS-EDTA solution (without  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ ), washed three times with ice-cold PBS, resuspended in adhesion buffer, and incubated with viruses as described above in the Examples section. As expected, all adenoviral serotypes tested are efficiently attached to and internalized into HeLa cells (Table III) (Figure 13). Adenoviruses serotypes 3, 5, 35, 41, but not 9, are efficiently attached to and internalized by 293 cells. In contrast, poor

attachment and internalization of most adenovirus serotypes are observed with CHO cells. The level of attachment on CHO is about 50-70 virus particles per cell for adenoviruses serotypes 5 and 41, 115 virus particles per cell for adenovirus type 3 and about 180 particles per cell for adenovirus serotypes 9 and 35. For further analysis, numbers >300 viral particles per cell are assumed as positive and <70 viral particles per cell as negative in terms of susceptibility of a particular cell line for efficient adenoviral transduction.

TABLE III

10 **Comparative analysis of attachment and internalization of Ad5 and Ad9 to cell lines, expressing different amounts of CAR and  $\alpha\beta$ integrins.**

| Cell line    | CAR expression | $\alpha\beta$ -integrin expression | Ad9 (attached/internalized) | Ad5 (attached/internalized) |
|--------------|----------------|------------------------------------|-----------------------------|-----------------------------|
| HeLa         | ++             | ++                                 | 426/370                     | 550/500                     |
| CHO          | -              | ++                                 | 300/300                     | 70/50                       |
| 293          | ++             | ++                                 | 20/20                       | 1950/1750                   |
| Y79          | +++            | -                                  | 190/140                     | 1200/1100                   |
| K562         | -              | +                                  | 320/230                     | 60/50                       |
| Erythrocytes | ?              | ?                                  | 420/-                       | 68/-                        |

15 Attachment and internalization of adenovirus serotypes 3, 5, 9, 35 and 41 into human CD34+ bone marrow cells and K562 erythroleukemia cell line:

Previous studies showed that the human erythroleukemia cell line K562 can be transduced with Ad5-based adenoviral vectors at very high MOIs. As shown in Figure 14, only about 60 viral particles per cell of adenovirus serotype 5 are attached to and even fewer particles are internalized into these cells at a MOI of 400. In contrast to Ad5, about 320 viral particles per cell of Ad9 and about 1500 viral particles per cell of Ad35 are

attached to and about two-thirds of them are internalized into K562 cells (Figure 14B). Human unstimulated CD34+-enriched bone marrow cells obtained from frozen stocks are incubated overnight in growth medium without cytokine stimulation. The next day, the number of viable cells is calculated. For attachment studies, cells are washed three times  
5 with ice-cold PBS, resuspended in adhesion buffer and incubated with adenoviruses. Among the adenoviral serotypes tested, only adenovirus particle of Ad9 (about 150 viral particles per cell) and Ad35 (about 320 viral particles per cell) are able to attach to unstimulated CD34+ cells on the level, compared to Ad 5 (only 60 viral particles per cell). Four-fifths of these virus particles are able to be internalized by the cells.  
10 Interestingly, upon stimulation of CD34+ cells with GM-CSF and EPO/TPO for two weeks, attachment and internalization of Ad9 viral particles are significantly increased (up to 300 particles per cell). At the same time, the transient stimulation of cells with GM-CSF for two days could not increase the level of viral attachment to the cells.

15 Based on the above finding that Ad35 serotype is able to attach and internalize into CD34+ cells most efficiently among several serotypes tested, serotype Ad35 was selected for further studies. As described in Appendix II, a chimeric vector (Ad5 GFP/F35) containing the short-shafted Ad35 fiber sequence in an Ad5 capsid was able to target a broad spectrum of CD34+ cells in a CAR/integrin independent manner.

20

#### *DISCUSSION*

In summary, from all the serotypes tested, Ad9, Ad3, and Ad35 demonstrated the most efficient attachment to and internalization with K562 and CD34+ cells. Based on  
25 adsorption/internalization data, Ad9 and Ad35 as representatives for subgroups D and B were selected for further tropism studies.

#### **D. Characterization of Ad vector replication in K562 and CD34+ cells.**

30 Comparative analysis of Ad5, and Ad9 and Ad34 to infect and to replicate in 293, K562 and CD34+ cells. The ability of the Ad9 fiber knob domain to recognize the same



primary receptor on the cell surface as Ad5 with comparable affinity was described earlier. Thus, the finding that Ad9 viral particles can only poorly attach to 293 cells is rather unexpected. In order to find out how the attachment and internalization data reflect the biological activity of adenoviruses of different serotypes, the stocks of Ad5, Ad9 and Ad35 are characterized in more detail by electron microscopy, plaque assay on 293 cells, and quantitative replication assay in K562 and CD34+ cells.

## METHODS

### Quantitative replication assay:

1x10<sup>5</sup> CD34+ or K562 cells were infected in 100µl of growth media with different MOIs of Ad5, 9, or 35 which had been amplified in 293 cells, expressing the XhoI DNA methyltransferase isoshizomer PaeR7 (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). After 2 hours of incubation at 37°C, the cells were centrifuged at 1000 x g for 5 min, the virus-containing medium was removed, the cells were resuspended in 100µl of fresh media, and then they were incubated at 37°C until harvesting. At 16 hours post-infection for K562 cells, or 36 h post-infection for CD34+ cells, 5 µg of pBS (Stratagene, La Jolla, CA) plasmid DNA was added as a carrier which could also be used as a loading control. Genomic DNA was extracted as described previously (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). One-fourth of purified cellular DNA (equivalent to 2.5 x 10<sup>4</sup> cells) was digested with HindIII, XhoI, or with HindIII and XhoI together at 37°C overnight and subsequently separated in a 1% agarose gel followed by Southern blot with chimeric Ad5/9 or Ad5/35 DNA probes. The chimeric probes, containing sequences of Ad5 and Ad9 (Ad 5/9) or Ad5 and Ad35 (Ad 5/35), were generated by a two-step PCR amplification using Pfu-Turbo DNA polymerase (Stratagene, La Jolla, CA) and viral DNA from purified particles as a template. The following primers were used for PCR (Ad5 sequences and nucleotide numbers are underlined): Ad5F1 – (nt: 32775-32805) 5'-GCC CAA GAA TAA AGA ATC GTT TGT GTT ATG-3'; Ad5R1 – (nt: 33651-33621) 5'-AGC TGG TCT AGA ATG GTG GTG GAT GGC GCC A-3'; chimeric Ad5/9F – (nt: 31150-31177, nt: 181-208 ) 5'-AAT

GGG TTT CAA GAG AGT CCC CCT GGA GTC CTG TCA CTC AAA CTA GCT  
GAC CCA -3'; chimeric Ad5/9R - (nt: 32805-32775, nt:1149-1113) 5'- CAT AAC  
ACA AAC GAT TCT TTA TTC TTG GGC TTC ATT CTT GGG CGA TAT AGG  
AAA AGG-3; chimeric Ad5/35F - (nt: 31150-31177, nt: 132-159) 5'-AAT GGG TTT  
5 CAA GAG AGT CCC CCT GGA GTT CTT ACT TTA AAA TGT TTA ACC CCA-3',  
chimeric Ad5/35R (nt: 32805-32775, nt: 991-958) 5'-CAT AAC ACA AAC GAT TCT  
TTA TTC TTG GGC ATT TTA GTT GTC GTC TTC TGT AAT GTA AG-3'.

Nucleotide numbers are given according to the sequences obtained from the NCBI  
GenBank (accession No. M73260 / M29978 for Ad5, X74659 for Ad9, and U10272 for  
10 Ad35). After the first amplification, the 968 bp-long Ad9, a 859 bp-long Ad35 DNA  
fragments corresponding to the fiber genes, and a 876 bp-long Ad5 fragment  
corresponding to the Ad5 E4 region (located immediately downstream of Ad5 fiber gene)  
were purified by agarose gel electrophoresis. To generate chimeric DNA probes,  
amplified Ad5 DNA was mixed with Ad9 or Ad35 fragments obtained during the first  
15 step of PCR, and subjected to a second PCR amplification using Ad5/9F or Ad5/35F  
primers and the Ad5R1 primer. The resulting Ad5/9 or Ad5/35 chimeric DNA fragments  
(see Fig. 15) were purified and their concentrations were measured  
spectrophotometrically. Corresponding chimeric DNA fragments were loaded as  
concentration standards on agarose gels or labeled with [<sup>32</sup>P]-dCTP and used as probes  
20 for Southern analysis. The number of viral genomes per DNA sample was calculated  
after quantitative Phospho-imager analysis. In preliminary experiments, no preferential  
hybridization of chimeric DNA probes to DNA of any particular viral serotype was  
detected.

## 25 RESULTS

### Replication of selected serotypes in K562 and CD34+ cells:

Adsorption/internalization studies do not ultimately prove viral transduction, a process  
30 often defined as gene transfer that allows for viral or heterologous gene expression in  
host cells. Intracellular trafficking, including endosomal lysis, transport to the nucleus,

and nuclear import of the viral genome, depends on structural capsid proteins and thus, varies between different serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065). We believed that analysis of viral gene expression would be a means to verify successful nuclear import of viral genomes and that this would be a good criterion for selection of serotype/s able to efficiently infect our target cells. To do this, we used a protocol, which allows for the detection of Ad replication in infected cells. Viral DNA synthesis can only occur after de novo expression of adenoviral early genes. We utilized a site-specific methylation strategy to monitor viral DNA replication within infected cells (Nelson, J., Kay, M.A. 1997. *Journal of Virology*. 71:8902-8907). Methylated Ad serotypes were produced by the addition of a methyl group onto the N6 position of the adenine base of Xho I sites, CTCGAG, during propagation of the viruses in 293 cells expressing the Xho I isoschizomer PaeR 7 methyltransferase (PMT) (Kwoh, T. J., et al. 1986. *Proc.Natl.Acad.Sci. USA*. 83:7713-7717) (293 PMTcells). Loss of methylation through viral replication restores Xho I cleavage and can be detected by Southern blots of Xho I-digested genomic DNA from infected cells.

Ad replication studies were performed in K562 and CD34+ cells with Ad9 and Ad35, in comparison to Ad5. For replication studies, the infectious titer (in pfu/ml) and genome titer (in genomes per ml) were determined (by plaque assay on 293 cells or by quantitative Southern blot, respectively) for methylated and unmethylated Ad5, Ad9, and Ad35 (Table 2). The ratio of pfu to genome titer was comparable for methylated and unmethylated virus demonstrating that DNA methylation had not altered transduction properties. About 85% of (Ad5, 9, and 35) virus used for infection was methylated as calculated based on the intensity of fragments specific for methylated and non-methylated viral DNA present in the viral load (Fig. 15). The numbers of genomes detected after adsorption (1 hour, 0°C) or internalization (2 hours 37°C) correlated well with studies shown in Fig. 12. Ad9 and Ad35 interacted more efficiently than Ad5 with K562 and CD34+ cells. Dose-dependent replication studies in K562 and CD34+ cells were performed with the same genome numbers of Ad5, 9, and 35 (Fig. 15). The replication rate was measured based on the ratio of methylated to demethylated viral

DNA after infection with different MOIs (2100, 420, and 105 genomes per cell). In K562 cells, efficient replication (100% conversion from methylated to unmethylated DNA) was detected for Ad5 at MOI  $\geq 2100$ , for Ad9 at MOI  $\geq 420$ , and for Ad35 at MOI  $\geq 105$ . This demonstrated that Ad35 transduced K562 cells with the highest efficiency.

5 In CD34+ cells, the replication rate was 100% for Ad5 and 31% for Ad9 after infection with MOI 420. Although methylated Ad35 viral DNA was present in CD34+ cells, viral replication was undetectable for Ad35. In summary, while viral replication studies in K562 cells confirmed data obtained for Ad5, 9, and 35 adsorption and internalization, there was a discrepancy between earlier results and the poor replication of Ad9 and, particularly, Ad35 in CD34+ cells. As outlined later, replication analysis in  
10 heterogeneous cell populations, like CD34+ cells, may not allow for definitive conclusions on tropism of a particular serotype.

Taking all the screening data together, Ad9 and Ad35 emerged as the variants with the strongest tropism for K562 and CD34+ cells. It is thought that Ad9 can bind to CAR,  
15 however, it preferentially uses  $\alpha_v$ -integrins for cell entry (Roelvink, P. W., et al. 1996. *J. Virology*. 70:7614-7621). This entry strategy may be not optimal for efficient infection of CD34+ cells as only less than 17% of them express  $\alpha_v$ -integrins (Fig. 10). Therefore, we decided to concentrate on Ad35 as a source for heterologous fiber to be used for  
20 construction of a chimeric vector based on an Ad5 backbone.

**TABLE IV**

**Results from the infectivity assay which determines the optical particle-to-PFU (OPU/PFU) ratio using 293 cells**

25

| Virus | OPU (A260)            | PFU                   | OPU/PFU ratio |
|-------|-----------------------|-----------------------|---------------|
| Ad5   | $1.4 \times 10^{12}$  | $1.06 \times 10^{11}$ | 13            |
| Ad9   | $4.61 \times 10^{11}$ | $2.6 \times 10^8$     | 1773          |

**DISCUSSION**

Viral replication studies in K562 cells confirmed the data obtained for Ad5, 9, and 35 adsorption and internalization. However, there was a discrepancy between the interaction data and the replication data in CD34+ cells where Ad9 replicated only poorly and no replication was seen for Ad35. Ad replication is only initiated upon the production of a critical threshold of early viral proteins, which in turn, is directly dependent on the number of viral genomes present in the nuclei of infected cells. Therefore, the outcome of replication studies may be affected by the rate of nuclear import of viral genomes, by the activity of viral promoters, and/or the intracellular stability of viral DNA/RNA. These parameters may vary, on one hand, between different subsets of CD34+, and/or, on the other hand, between different Ad serotypes. In conclusion, the viral replication analyses performed with different Ad serotypes in CD34+ cells may not predict the actual transduction properties of chimeric vectors based on Ad5 backbone. This implies that attempts to produce gene transfer vectors based on Ad genomes other than Ad5 should be exercised with caution.

Recently, an Ad serotype screening strategy was used to identify variants with tropism for primary fetal rat CNS cortex cells or human umbilical vein endothelial cells. The optimal serotype (Ad17) was selected based on immunohistochemistry for hexon production 48 hours after infection (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540). However, this approach is problematic because, at least in our hands, antibodies developed against Ad5 hexon did not cross-react with other serotypes. Also, hexon is expressed only after onset of replication. As outlined above, the kinetics of intracellular trafficking, viral gene expression, and replication significantly vary between serotypes (Defer, C., et al., P. 1990. *J. Virology*. 64:3661-3673; Miyazawa, et al. 1999. *J. Virology*. 73:6056-6065).

In addition to being the most efficient serotype in terms of interaction with CD34+ cells, Ad35 is also interesting because it interacts with receptor/s different from the Ad5 and Ad3. Ad35 and Ad5GFP/F35 attachment was not inhibited by Ad5 or anti-CAR antibodies suggesting that Ad35 binding was CAR independent. First, Ad5 did not compete with Ad35 and Ad5GFP/F35 during internalization and infection indicating that

$\alpha_5\beta_3/5$  integrins are not involved in viral entry. Second, function-blocking antibodies against  $\alpha_v$ -integrins did not compete with Ad35 and Ad5GFP/F35 for internalization into K562 cells, whereas these antibodies did inhibit Ad5 internalization. And third, in contrast to Ad5 based vectors, GFP expression after infection with Ad5GFP/F35 was not restricted to  $\alpha_v$ -integrin-expressing CD34+ cells. From these facts, we conclude that infection with Ad35 and the chimeric Ad5GFP/F35 vector does not involve  $\alpha_v$ -integrins. In this context, the presence or absence of RGD motifs within Ad35 penton base remains to be determined by sequencing the corresponding genome region. Cross-competition assays demonstrated that Ad35 and Ad5GFP/F35 bind to a receptor that is different from the Ad3 receptor. Although Ad3 and 35 belong to the same subgroup, they have been divided into two DNA homology clusters, B1 and B2; the amino acids composing their fibers are only 60% homologous. Furthermore, the target tissues for both viruses are different; Ad3 can cause acute respiratory infections, whereas Ad35 is associated with kidney infection (Horwitz, M. S. 1996. *Adenoviruses*, p. 2149-2171. In B. N. Fields, Knipe, D.M., Howley, P.M. (ed.), *Virology*, vol. 2. Lippincott-Raven Publishers Inc., Philadelphia). Therefore, it was not surprising to see that Ad3 and Ad35 recognize different receptors.

In conclusion, Ad35 and the chimeric vector enter the cells by a CAR- and  $\alpha_v$ -integrin independent pathway. We believe that Ad35 and the chimeric vector binds primarily to its fiber receptor and that this interaction is sufficient to trigger internalization. On the other hand, Ad35 internalization may involve cellular proteins other than  $\alpha_v$ -integrins. These membrane proteins can overlap with those for Ad3 internalization and represent  $\beta_2$  integrins, which protrude more from the cell surface than  $\alpha_v$ -integrins (Huang, S., et al. 1996. *J. Virology*. 70:4502-4508).

According to EM studies of negative contrast-stained adenoviral suspensions, the percentage of deficient particles for all adenoviral serotypes tested does not exceed 5%. However, plaque assays reveal that the ability to form plaques in 293 cells is significantly different for tested serotypes. The optical particle-to-PFU (OPU/PFU) ratio obtained is 13 for Ad5, which is in good agreement with the previously estimated ratio for this

adenoviral serotype. Importantly, this ratio is about three times higher for adenovirus serotype 35 and more than 150-fold higher for adenovirus serotype 9. Furthermore, quantitative Southern blot using chimeric Ad5/9 and Ad5/35 DNA probes is used to determine the ratio between the genome and transducing titer. This study confirms the data obtained by plaque assay. Quantitative replication assay of these adenoviruses in K562 and CD34+ cells also confirms the ability of Ad9 and Ad34 to more efficiently attach to these cell types. The replication of viral genomes is observed for Ad9 and Ad34 at lower MOIs of infection, compared to Ad5. In conclusion, the data obtained for different serotypes in attachment and internalization are in good agreement with the infectivity data in target cells.

**E. Attachment and internalization of different adenoviral serotypes into primary dendritic cells, JAWSII, MCF-7 and REVC cells.**

As a proof of principle, the serotype screening strategy can be employed for other important target cells which are refractory to Ad5 infection.

**RESULTS**

RECV cells are endothelial cells which have to be targeted for approaches that are aimed to gene therapy of restenosis, atherosclerosis, inflammation etc. MCF-7 cells are breast cancer cells isolated from liver metastases which are important targets for tumor gene therapy. The human adenovirus serotypes 3, 5, 9, 35 and 41 are tested to see whether they can attach to and can be internalized by mouse primary dendritic cells, JAWSII cells, MCF-7-human breast cancer cells and REVC endothelial cells. None of the adenoviral serotype tested can efficiently attach to primary dendritic cells. Adenovirus serotype 3 is able to efficiently attach to REVC endothelial cells (about 400 virus particles per cell are attached and about 300 are internalized). In comparison, only 50 Ad5 particles are able to attach to and even fewer are internalized in these REVC. The human breast cancer cells (MCF-7) are previously shown to be refractory to Ad5

infection at low MOIs. However, Ad3 and more efficiently, Ad35 attach to and internalize into MCF-7 cells.

### *DISCUSSION*

5

The data presented herein indicate that different human adenovirus serotypes recognize different cellular receptors and can therefore infect cell types that are refractory to Ad5 infection. There are adenoviral serotypes that can more efficiently attach and internalize than Ad5 for human CD34+ cells, REVC, K562 and MCF-7 cells. This finding provides  
10 a basis for the construction of chimeric adenoviral vectors which are Ad5 vectors containing receptor ligands derived from other serotypes.

#### **F. Infection studies on primary human bone marrow cells.**

15 Since established erythroleukemic cell lines do not represent an adequate model for the ultimate hematopoietic stem cell that has to be targeted in patients in order to achieve long-term reconstitution with genetically modified cells, normal primary human bone marrow cells are used for the initial infection/retargeting studies.

### 20 *RESULTS*

In a first set of tropism studies with different Ad serotypes, whole bone marrow cell suspensions can be used without preselection. This is advantageous because the tropism of various adenovirus serotypes or genetically retargeted vectors can be analyzed on a  
25 broad spectrum of progenitor subpopulations representing myeloid, erythroid, megakaryocytic, lymphoid, dendritic, and monocytic lineages. For short term (< 5 hours) infection studies, bone marrow suspensions can be cultured in IMDM supplemented with 10% FCS,  $\beta$ -mercaptoethanol, and 10u/ml IL-3 for ensuring cell viability.

30



Mononucleated cell assays:

Mononucleated bone marrow cells can be incubated with MOI 1, 10, 100, or 1000 pfu/cell of the various adenovirus types for a short time. Paraffin sections or cytopins of infected bone marrow cells can be analyzed for nuclear-localized, labeled viral DNA. BrdU labeling can be visualized by immunofluorescence with anti-BrdU antibodies; <sup>32</sup>P-tagged viral DNA can be detected by incubation with photo-emulsion. In addition, the same cell material can be analyzed for morphology after specific histo-staining (e.g. Wright-, Hemo3 staining). If required, commercially available antibodies can be used to specific cell surface markers conjugated directly to different fluorochromes (FITC (green), TRIT., RPE, (red), RPE-Cy5, AMCA (blue)) to completely characterize infected bone marrow subpopulations. Colocalization of BrdU-labeled viral DNA (e.g. as FITC signal) with membrane markers signifying infection of specific cell types can be demonstrated; for example, potential stem cells/early progenitors (CD34<sup>+</sup>, CD38<sup>-</sup>), megakaryocytes (CD41a<sup>+</sup>), erythroid cells (glycophorin A<sup>+</sup>), dendritic cells (CD1a<sup>+</sup>), monocytes (CD14<sup>+</sup>), or myeloid cells (CD15<sup>+</sup>), etc. The morphological analysis of infected bone marrow subsets gives a first information whether specific adenovirus serotypes can target primitive cell types.

20 *DISCUSSION*

Since the different wild-type adenoviruses do not express a uniform marker gene and do not integrate and since detection of tagged viral DNA cannot be done on live cells, it is not possible, at this point, to characterize infected cells for clonogenic or repopulation capacities. Therefore, adenovirus serotypes for retargeting studies are selected, based on their ability to infect in vitro purified CD34<sup>+</sup> cells at low MOIs. This subset of bone marrow cells is known to contain long-term reconstituting cells. Infection studies with different adenovirus serotypes can be repeated on purified CD34<sup>+</sup> cells (cultured in IMDM +10% FCS, β-mercaptoethanol, and 10 units/ml IL-3) as described above. Purification of CD34<sup>+</sup> cells can be performed by direct immunoadherence on anti-CD34 monoclonal antibody-coated plates or on MiniMacs columns as described by

Papayannopoulou (Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69; Papayannopoulou, T. et al., 1993, *Blood*, 81:229). The purity of isolated CD34+ cells ranges routinely from 80-95%. Analogous infection studies can be repeated with selected adenovirus types on CD34+/CD38- subsets.

5

To confirm productive infection purified CD34+ cells can be infected with selected (methylase-tagged) serotypes and analyze viral DNA replication. Cultures of purified human bone marrow CD34+ cells can be used for the transduction and integration studies as a model for HSCs.

10

It was recently demonstrated that HSC activity does exist in CD34-negative human bone marrow subsets (Bathia, M. et al., 1998, *Nature Medicine*, 4:1038-45; Osawa, M., et al., 1996, *Science*, 273:242-5; Goodell, M. et al., 1997, *Nature Medicine*, 3:1337-45; Zanjani, E. D. et al., 1998, *Exp. Hematology*, 26:353-60). Lin<sup>-</sup>CD34<sup>-</sup>CD38<sup>-</sup> cells can be tested in the retargeting and transduction studies in combination with repopulation assays in SCID-NOD mice.

15

#### G. Cloning and insertion of the fiber gene.

#### 20 *METHODS*

##### PCR-cloning of the corresponding fiber gene and insertion into Ad5 based shuttle plasmids instead of the endogenous AD5 fiber:

25 One or several adenoviruses with tropism to CD34+ or other HSC containing population is selected for further studies described herein. The complete coding region for fiber varies between 1-2kb, depending on the virus type. The fiber encoding sequences can be obtained by PCR with Pfu polymerase from viral DNA isolated from purified particles of the selected virus types. The corresponding primers can be designed based on the fiber sequences available from the EMBL gene bank. The PCR products are cloned as PacI-Ball  
30 fragment into pCD4 (Figure 10), a shuttle vector for recombination of RecA+ E.

coli. In pCD4, the heterologous fiber gene is flanked on both sides with Ad5 sequences, which are homologous to regions directly adjacent to the fiber reading frame in Ad5. As an Ad5 (shuttle vector) derived template for recombination, pCD1, a pBHG 10 (Microbix, Toronto, Canada) derivative can be used. The recombination procedure is performed according to a protocol routinely used for recombinant adenovirus generation (Chartier, C., et al., 1996, *J. of Virology*, 70, 4805-4810). Routinely, 90% of the resulting plasmids are accurately recombined. The junctions between the heterologous fiber (X) and Ad5 sequences can be sequenced to confirm the accuracy of recombination. The resulting plasmid is named pAd5fiberX (pAd5<sup>fx</sup>). The resulting product is used to generate pAd5<sup>fx</sup>-based Ad.AAV containing the heterologous fiber gene.

Construction of chimeric Ad vectors:

For transduction studies, two Ad vectors were constructed: Ad5GFP and Ad5GFP/F35, containing a chimeric Ad5/35 fiber gene. Both adenoviral vectors contained a 2.3kb, CMV promoter driven EGFP gene [derived from pEGFP-1, (Clontech, Palo Alto, CA)] inserted into the E3 region of Ad5. The EGFP expression cassette was cloned between Ad5 sequences 25,191-28,191 and 30,818-32,507 into a shuttle plasmid, which contained the E3 deletion described for pBHG10 (Microbix, Toronto, Canada). The resulting plasmid was named pAdGFP. For the chimeric vector, the Ad5 fiber gene in pAdGFP was substituted by an Ad5/35 chimeric fiber gene generated by the two-step PCR protocol outlined above. In the first PCR step, three DNA fragments corresponding to i) the Ad5 fiber 5'-nontranslated region and the first 132 bp of the fiber tail domain (nt 30,818-31,174), ii) the Ad35 shaft and knob domains (nt 132-991), and iii) the Ad5 E4 region including the Ad5 fiber polyadenylation signal (nt 32,775-33,651) were amplified by Pfu-Turbo DNA polymerase. The following primers were used: for the Ad5 tail, Ad5F-2 (nt 30,798-30,825) 5'-CGC GAT ATC GAT TGG ATC CAT TAA CTA-3' and Ad5R-2 (nt 31,174-31,153) 5'-CAG GGG GAC TCT CTT GAA ACC CAT T-3'; for the Ad35 shaft and knob, primers Ad5/35F and Ad5/35R (see above); for the Ad5E4 and polyA, primers Ad5F-1 and Ad5R-1 (see above). After 10 PCR cycles, the products were purified by agarose gel electrophoresis, combined, and then subjected to a second PCR

with primers Ad5F-2 and Ad5R-1. The resulting 2115 bp-long chimeric fiber gene contained the Ad5 tail and the Ad35 shaft and knob domains. This product was used as a substitute for the Sall/XbaI Ad5 fiber gene containing fragment in pAdGFP. The resulting plasmid was named pAdGFP/F35. To generate full-length E1/E3 vector genomes, pAdGFP and pAdGFP/F35 were inserted in pAdHM4 (Mizuguchi, H., Kay, M.A. 1998. *Human Gene Therapy*. 9:2577-2583) by recombination in *E.coli* (Chartier, C., E. et al. 1996. *Journal of Virology*. 70:4805-4810). To do this, the RecA+ *E.coli* strain BJ5183 was co-transformed with pAdHM4 linearized by SrfI mixed with the XbaI fragments containing the GFP genes, the Ad5 or Ad5/35 fiber genes, and the Ad5 homology regions. The resulting recombinants were analyzed by restriction analysis. Correct recombinants were amplified in *E.coli* HB101 and purified by double CsCl gradient banding. The plasmids were named pAd5GFP and pAd5GFP/F35. The correct structure of the Ad5/35 chimeric fiber gene was confirmed by endonuclease digestion and sequencing part of pAd5GFP/F35. To produce the corresponding viruses, pAd5GFP and pAd5GFP/F35 were digested with PacI to release the viral genomes and transfected onto 293 cells as described (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960). Plaques developed 7 to 10 days post-transfection in overlaid cultures. Recombinant viruses were propagated in 293 cells and purified by standard methods described elsewhere (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

#### Hemagglutination assay:

Twenty-five microliters of serial dilutions of Ad5, Ad35, or chimeric Ad5GFP/F35 virions in McIlvaine-NaCl buffer (0.1 M citric acid, 0.2 M Na<sub>2</sub>HPO<sub>4</sub> [pH 7.2], diluted 1:50 with 0.87% NaCl) were loaded onto 96 well plates. To each dilution, 25 µl of a 1% suspension of monkey erythrocytes (in McIlvaine-NaCl buffer) was added. The sedimentation pattern was determined after incubation for 1 hour at 37°C. All tests were performed in quadruplicates in at least two independent experiments.

Southern blot:

Extraction of genomic DNA, labeling of DNA fragments and hybridization were performed as described earlier (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960).

RESULTSConstruction/Characterization of chimeric fiber:

Previously, it has been shown that exchanging the fiber knob was sufficient to alter the tropism of chimeric Ad vectors (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V., et al. 1998. *J. Virology*. 72:1844-1852; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). As outlined above, the length of the fiber shaft may critically determine the entry strategy of a particular serotype. Therefore, we decided to replace not only the Ad5 fiber knob but also the shaft. The chimeric Ad5/35 fiber contained the Ad5 tail (amino acid: 1-44) necessary for interaction with the Ad5 penton base linked to 279 amino acids from Ad35 including the shaft with 7  $\beta$ -sheets and the knob (Fig. 16A). The endogenous Ad5 fiber polyA signal was used to terminate transcription of the chimeric fiber RNA. The combination of the Ad5 capsid including the RGD motif containing penton base with a short-shafted fiber could be risky because the natural distance between the fiber knob and the RGD motifs was disturbed. The Ad5 fiber was substituted by the chimeric fiber sequences based on an E1/E3 deleted Ad vector. This vector carried a CMV promoter-GFP reporter gene cassette inserted into the E3 region. The corresponding chimeric virus (Ad5GFP/F35) was produced in 293 cells at a titer of  $>2 \times 10^{12}$  genomes per ml. For comparison, an E1/E3 deleted Ad vector containing the original Ad5 fiber gene and the GFP expression cassette was generated (Ad5GFP). The titer and the ratio of physical to infectious particles was similar between Ad5GFP and Ad5GFP/F35 indicating that the fiber modification did not significantly alter the stability and/or growth properties of the chimeric vector. The correctness of the fiber modification was confirmed by restriction analysis of the Ad5GFP/F35 viral genome followed by

- Southern blot hybridization (Fig. 16B), direct sequencing of the fiber-coding region, and a functional test for hemagglutination (HA) of monkey erythrocytes. The agglutination of erythrocytes is fiber knob-mediated; it is known that Ad5 does not agglutinate monkey erythrocytes whereas Ad35 efficiently does (Pring-Akerblom, P., et al. 1998. *J. Virology*. 72:2297-2304). In HA tests, Ad5GFP/F35 agglutinated monkey erythrocytes with the same efficiency as Ad35 at dilutions of up to 1:512. In contrast, no hemagglutination was observed with equivalent Ad5 dilutions. This clearly confirmed the functional activity of the chimeric Ad5/35 fiber incorporated into Ad5 capsid.
- 10 Generation of chimeric adenoviral vectors (Ad.AAV<sup>fk</sup>) with heterologous fiber molecules: Adenoviruses with chimeric Ad5-Ad3 fiber are viable and can be produced at high titers (Krasnykh, V., et al., 1996, *J. of Virology*, 70, 6839-6846; Stevenson, S. C. et al., 1997, *J. Virology*, 71:4782-90). In order to test whether the fiber substitution described herein affects production or stability of adenoviruses, two E1-deleted first-  
15 generation, adenoviral vectors are produced with the AAV- $\beta$ gal cassette in 293 cells using standard protocols. The vector is generated by recombination of pAd.AAV-BG (Fig. 17) with pCD1 (containing the endogenous Ad5 fiber); the other vector (with heterologous fiber) is the recombination product of pAd.AAV  $\beta$ Gal and pAd5fiberX (pAd5<sup>fk</sup>). Virus from single plaques is amplified on 293 cells. The production yield per  
20 293 cell can be determined by plaque-titering of 293 cell Lysates. It is anticipated that the fiber modification will not critically affect the stability of chimeric vectors. Finally, bone marrow cells can be infected with the retargeted vectors. Two days after infection, live-cell cytometry is performed for  $\beta$ -gal expression using as substrate Fluorescein di- $\beta$ -D-Galactopyranoside (FDG) (Cantwell, M.J. et al., 1996 *Blood* 88, 4676-4683; Neering, S. et al., 1996, *Blood*, 88:1147-55; Fiering, S. N. et al., 1991, *Cytometry*, 12:291; Mohler, W. et al., 1996, *PNAS*, 93:57) and the infected cells are characterized for morphology and surface markers. Before and during infection, bone marrow cells can be cultured in IMDM/FCS supplemented with thrombopoietin (Tpo), which supports the survival of HSC (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996,  
25 *Experimental Hematology*, 24:660-69). Alternatively, retargeted vectors can be  
30

generated with the AAV-GFP (green fluorescence protein) cassette and perform FACS analysis on transduced cells based on GFP and surface marker expression.

#### **H. Competition studies of chimeric fiber protein Ad5/35.**

5

##### **Competition studies:**

Cross-competition studies between Ad5, 35, and Ad5GFP/F35 (Fig. 18) for binding and internalization were performed in order to investigate in more detail the pathways which are used by the chimeric vector to infect target cells. Wild-type Ad35 and the chimeric vector Ad5GFP/F35 could recognize the same primary receptor as they competed with each other for the attachment to K562 cells (Fig. 19A, upper panel). This primary receptor is different from that used by Ad5, since neither Ad5 viral particles nor anti-CAR monoclonal antibodies (Fig. 19B, upper panel) were able to abrogate Ad35 or Ad5GFP/F35 binding. In competition studies for internalization, Ad35 and Ad5GFP/F35 competed with each other with equal efficiency. Ad5 and anti- $\alpha_v$ -integrin monoclonal antibodies (L230) (Figs. 19C, D; lower panel) did not inhibit internalization of Ad35 or the chimeric virus. To consolidate this data, K562 cells were infected with Ad5GFP and Ad5GFP/F35 after prior incubation of cells with anti-CAR or anti- $\alpha_v$ -integrins monoclonal antibodies followed by analysis of GFP-expressing cells. The transduction data mirror the results obtained in adsorption/internalization studies. In summary, this demonstrated that Ad35 and Ad5GFP/F35 use a CAR and  $\alpha_v$ -integrin-independent pathway for infection of K562 cells; the structural elements which account for these specific properties are located within the Ad35 fiber and can be transplanted into Ad5 by fiber substitution.

Ad3 can efficiently interact with K562 cells (Fig. 12), although Ad3 and Ad35 belong to the same subgroup (B), the homology between amino acid sequences of their fibers is only about 60%. Therefore, we decided to test whether Ad3 could compete with Ad35 and Ad5GFP/F35 for attachment and internalization (Fig. 20). These studies demonstrated that Ad35 binding was not inhibited by Ad3 indicating the use of different

30

receptors. Interestingly, Ad3 slightly inhibited attachment of Ad5GFP/F35 (Fig. 20A, left panel). In addition to binding to the receptor common for the Ad35 and Ad5GFP/F35 fiber, the chimeric capsid (e.g. the Ad5 penton RGD motifs) may also interact with a second cellular receptor that overlaps with elements involved in Ad3 binding. In cross-competition for internalization, pre-incubation of cells at 37°C with Ad35 and with chimeric virus significantly decreased internalization of [<sup>3</sup>H]-labeled Ad3 (Fig. 20D, right panel). In the reverse experiment, Ad3 as competitor decreased the level of internalization by 30% for both, Ad35 and the chimeric virus (Fig. 20B, right panel). As expected, Ad5 and Ad3 did not compete for adsorption or internalization. As shown before (Fig. 19B), anti-CAR and anti- $\alpha_v$ -integrin antibodies did not block Ad3 interaction with K652 cells. In summary, we concluded that Ad35 and Ad5GFP/F35 bind to receptor/s different from that of Ad3, although they can use common structural elements for internalization, which are different from  $\alpha_v$ -integrins.

15 Infection studies with chimeric virus:

It is established that Ad5GFP/F35 infected K562 cells by a CAR and  $\alpha_v$ -independent pathway. It is possible that this property allows for efficient transduction of non-cycling CD34+ cells, which express scarcely CAR and  $\alpha_v$ -integrins. To test this, the transduction properties of Ad5GFP and Ad5GFP/F35 vectors were analyzed on CD34+ cells, K562, and HeLa cells. Fig. 21 shows the percentage of transduced, GFP expressing cells depending on the MOI used for infection. Nearly 100% of HeLa cells were transduced with Ad5GFP and Ad5GFP/F35 at MOIs of  $\geq 25$ . More than 95% of the K562 cells were transduced with Ad5GFP/F35 at MOIs of  $\geq 100$ , whereas the transduction rate was significantly lower with Ad5 where it increased with the MOI reaching a plateau at ~70% GFP-positive cells after infection with an MOI of 400. Transduction of CD34+ cells was about three fold more efficient with Ad5GFP/F35 than with Ad5GFP at all MOIs analyzed. Interestingly, at higher MOIs, the transduction rate did not rise proportionally with the viral dose and soon reached a plateau indicating that in both cases only specific subset/s of CD34+ cells were permissive to infection. In order to characterize in more detail these specific, permissive subset/s, additional transduction



studies were performed. First, the percentage of GFP expressing cells was determined in CD34+ fractions that were stained for  $\alpha_v$ -integrins or CARs (Fig. 22). The low number of CAR positive CD34+ cells complicated accurate co-labeling studies, and there was no correlation between CAR expression and the proportion of transduced cells among CD34+ cells infected with Ad5GFP or Ad5GFP/F35. Interestingly, for Ad5GFP, 65% of all GFP expressing cells were positive for  $\alpha_v$ -integrins, whereas less than 22% of GFP positive cells infected with the chimeric virus stained positive for  $\alpha_v$ -integrin expression. While only 17% of the whole CD34+ population expressed GFP after Ad5GFP infection, the percentage of GFP-expressing cells in the CD34+/ $\alpha_v$ -integrins positive fraction was 50%. This indicates that Ad5GFP vector-mediated GFP expression was preferentially localized to  $\alpha_v$ -integrin positive CD34+ subsets, whereas after infection with the Ad5GFP/F35 vector, GFP was expressed in a broader spectrum of CD34+ cells with most of them being  $\alpha_v$ -integrin-negative.

Next, transduced cells were simultaneously analyzed for GFP as well as for CD34 and CD117 markers. As mentioned before, only about 90% of all cells used in our analysis were positive for CD34 at the time of infection, hence the multiparameter analysis for CD34 and GFP. A population of CD34+ cells is extraordinarily heterogeneous in morphology and stem cell capacity. The subpopulation of CD34+ and CD117+ cells resembles very primitive hematopoietic cells (*Ikuta, K., Weissman, I.L. 1992 Proc. Natl. Acad. Sci. USA. 89:1502-1506; Simmons, P. J., et al. 1994. Expl. Hematology. 22:157-165*). Fig. 23 summarizes the analyses of GFP expression in correlation with these specific stem cell markers. While 54% of cells infected with chimeric vector were positive for GFP and CD34+, only 25% of cells infected with Ad5GFP expressed the transgene and CD34+ marker (Fig. 23A, lower panel). More importantly, based on GFP expression, the chimeric virus transduced 80% of c-kit positive cells, whereas the Ad5-based vector transduced only 36% (Fig. 23A, middle panel). In an additional experiment, CD34+ cells were sorted for CD117 expression prior to infection with Ad5GFP or Ad5GFP/F35 and, 24 hours post-infection, GFP expression was analyzed in this specific fraction (Fig. 23B). This analysis revealed that the chimeric vectors transduced 4 fold more CD34+/CD117+ than the Ad5GFP vector.

In conclusion, these results demonstrated that the chimeric Ad5GFP/F35 vector was clearly superior to the Ad5GFP vector in targeting and transduction of CD34+ cells. Furthermore, the data suggest that the spectrum of CD34+ cell subsets permissive for Ad infection was significantly different for the chimeric vector than for the Ad5 vector.

Analysis of viral genomes within CD34+ cells infected with the Ad5 and chimeric vectors:

So far, the transduction rate of CD34+ cells was measured based on GFP expression after infection with Ad5GFP and Ad5GFP/F35. Considering the extraordinary heterogeneity of CD34+ cells in morphological and functional parameters, GFP may not be expressed in all cell types that were efficiently infected. Reasons for this include that the CMV promoter may not be active in all cell types or that the regulation of transgene expression could differ between subsets on a post-transcriptional or post-translational level. To test this, we quantified the number of intracellular (transduced) viral genomes within GFP positive and GFP negative fractions of CD34+ cells infected with Ad5GFP and Ad5GFP/F35. To do this, twenty-four hours after infection, CD34+ cells were sorted for GFP positive and GFP negative fractions, which were subsequently used to isolate genomic DNA together with transduced viral DNA. The number of viral genomes was determined by quantitative Southern blot as described for Fig. 15. Per GFP-positive CD34+ cell, about 270 copies of the Ad5GFP/F35 viral genome were detected. Interestingly, a remarkable 200 copies of the Ad5GFP/F35 viral genome were found per GFP-negative CD34+ cell (Fig. 24A and 25). This demonstrated that not all infected cells expressed GFP and implies that the actual transduction rate was higher than 54% (GFP-positive cells). We concluded that the CMV promoter was not active in all transduced CD34+ subsets. No Ad5GFP vector specific signal was detected within infected CD34+ (GFP positive or negative) fractions by Southern blot which had a detection limit of 14 viral genomes per cell. From this, we can conclude that the vector DNA concentration per transduced cell was at least 20 times higher for Ad5GFP/F35 than for Ad5GFP.

Ad5GFP DNA was only detectable in DNA samples from infected CD34+ cells by Southern blot after prior PCR amplification with vector specific primers (Fig. 24B and 25). This indicates that the replication deficient Ad5 vector is present but at a very low copy number, which may be limited by intracellular genome stability. Using the PCR-  
5 Southern detection method, Ad5 vector DNA was also detected in GFP negative cells, supporting that the CMV promoter may not have been the optimal choice for transduction studies. It is notable that studies by others on viral genome analyses after infection of CD34+ cells with Ad5 vectors were performed only after prior PCR amplification (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Neering, S. J., et al. 1996..  
10 *Blood*. 88:1147-1155).

### DISCUSSION

The chimeric Ad5GFP/F35 vector has binding and internalization properties similar to  
15 Ad35. Therefore, the fiber substitution was sufficient to swap cell tropism from Ad5 to Ad35. The Ad5GFP/F35 capsid chimera contained the short-shafted Ad35 fiber incorporated into an Ad5 capsid, instead of the naturally occurring long-shafted Ad5 fiber. During Ad5 infection, interaction between the penton base and integrins is required to induce viral internalization. For this interaction, the length of fiber shaft and  
20 the precise spatial arrangement of knob and RGD motifs are critical for the virus entry strategy. The natural spatial arrangement is disturbed when short-shafted heterologous fibers are inserted into the Ad5 capsid. Interestingly, the Ad5/35 capsid chimera allows for efficient infection, suggesting that the protruding RGD motives in the Ad5 penton base do not affect the interaction with the primary Ad35 receptor. So far, most of the  
25 chimeric viruses were generated by substituting only the Ad5 knob while maintaining the long Ad5 fiber shaft (Chillon, M., et al. 1999. *J. Virology*. 73:2537-2540; Krasnykh, V. N., et al. 1996. *J. Virology*. 70:6839-6846; Stevenson, S. C., et al. 1995. *J. Virology*. 69:2850-2857; Stevenson, S. C., et al. 1997. *J. Virology*. 71:4782-4790). The exception was an Ad5/7 chimeric virus (Gall, J., et al. 1996. *J. Virology*. 70:2116-2123), where the  
30 whole Ad5 fiber was substituted by the short-shafted Ad7 fiber. However, similar to the parental Ad5, the Ad5/7 chimera still required  $\alpha_v$ -integrins for infection.

This Ad5GFP/F35 chimera is the first demonstration that despite the presence of RGD motifs within the Ad5 penton, the chimeric virus uses cell entry pathways determined primarily by the receptor specificity of the short-shafted heterologous fiber. This does not  
5 exclude that interaction with a secondary receptor may increase binding affinity. The latter is supported by the observation that Ad35 and Ad5GFP/F35 slightly differed in their ability to compete with Ad5 or Ad3 for binding. It is possible that Ad5/35 attachment involves, in addition to the high affinity fiber binding, interaction between Ad5 capsid proteins (e.g. RGD motifs) and secondary receptor/s that overlap with those  
10 used by Ad3 and Ad5.

This data indicate that infection with Ad5-based vectors is restricted to a specific subset of CD34+ cells. The percentage of GFP expressing cells after Ad5GFP infection of CD34+ cells reached a plateau at MOIs higher than 100 indicating that only a limited  
15 fraction of CD34+ cells was permissive to Ad5. Also, strong replication of wild type Ad5 in infected CD34+ cells may be the result of preferential transduction of a specific subpopulation of CD34+ resulting in a expression of early viral genes at a level sufficient to initiate viral replication. The presence of a specific subpopulation of CD34+ cells permissive to Ad5-vector infection was suggested by others (Byk, T., et al. 1998. *Human*  
20 *Gene Therapy*. 9:2493-2502; Neering, S. J., et al. 1996.. *Blood*. 88:1147-1155). In the present report, we further characterized this subpopulation and demonstrated that Ad5-based vectors preferentially infected  $\alpha_v$ -integrin positive CD34+ cells. Integrins (including  $\alpha_v$ ) are thought to be important for homing and trafficking of transplanted hematopoietic cells, however little is known about the correlation between  $\alpha_v$ -integrin  
25 expression and the differentiation status of hematopoietic cells (Papayannopoulou, T., Craddock, C. 1997. *Acta Haematol.* 97:97-104; Roy, V., Verfaillie, C.M. 1999. *Exp. Hematol.* 27:302-312). There was no clear correlation between CAR and GFP expression suggesting that Ad5GFP may be able to use another membrane protein as a primary receptor. Alternatively, Ad5GFP transduction observed at an MOI of 200-400 could be  
30 the result of direct interaction between virus and  $\alpha_v$ -integrins triggering internalization, which may be the preferred pathway in the absence of CAR (Legrand, V., et al. 1999. *J.*

*Virology*. 73:907-919). Importantly, infection with the chimeric Ad5GFP/F35 vector was not restricted to the  $\alpha_v$ -positive CD34+ subpopulation.

Among CD34+ cells, the subpopulation of CD34+ and CD117+ cells resembles very primitive hematopoietic cells (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*. 89:1502-1506; Simmons, P. J., et al. 1994. *Expl. Hematology*. 22:157-165). The receptor for stem cell factor, CD117 (c-kit) belongs to a tyrosine kinase family. It was previously shown that c-kit+, CD34+ cord blood cells contain a high fraction (16%) of hematopoietic progenitors (Neu, S., et al. 1996. *Leukemia Research*. 20:960-971). Early in ontogeny 34+/CD117+ cells have long-term repopulating activity (Sanchez, M. J., et al. 1996. *Immunity*. 5:513-525). An average of 50-60% of CD34+ cells are reported to be CD117 positive (Ikuta, K., Weissman, I.L. 1992 *Proc. Natl. Acad. Sci. USA*. 89:1502-1506; Neu, S., et al. 1996. *Leukemia Research*. 20:960-971; Simmons, P. J., et al. 1994. *Expl. Hematology*. 22:157-165). In our studies, the chimeric vector expressed GFP in 54% CD34+ cells and 80% of CD34+/c-kit+ cells. The actual viral transduction rate could be even higher because transduced Ad5GFP/F35 vector DNA was also found in GFP-negative fractions of infected cells. This indicates that the CMV promoter used to drive GFP expression in our vectors was not active in all transduced cells. We selected the CMV promoter for transgene expression based on published data demonstrating that PGK and CMV promoters allowed for efficient transgene expression in CD34 cells whereas the HTLV-I and RSV promoter were almost inactive (Byk, T., et al. 1998. *Human Gene Therapy*. 9:2493-2502; Case, S. S., et al. 1999. *Proc. Natl. Acad. Sci. USA*. 96:2988-2993). On the other hand, studies by Watanabe et al. (Watanabe, T., et al. 1996. *Blood*. 87:5032-5039) suggest that the CMV promoter is not active or rapidly silenced in certain CD34+ subsets. Our data underscore this observation. Considering retroviral transduction studies, the retroviral MLV promoter may have been a better candidate for transduction studies in hematopoietic cells (Bregni, M., et al. 1998. *Gene Therapy*. 5:465-472).

After having demonstrated that the Ad5GFP/F35 vector efficiently transduced cells carrying stem cell specific markers, the next logical step would be to perform colony

assays with pre-sorted GFP positive/negative cells. However, this assay is complicated by the fact that infection with first generation Ad vectors is cytotoxic and affects the formation and growth of progenitor colonies in MC-cultures (Mitani, K., et al. 1994. *Human Gene Therapy*. 5:941-948; Watanabe, T., et al. 1996. *Blood*. 87:5032-5039). This side effect is caused by the expression of Ad proteins within transduced cells (Lieber, A., C.-Y. et al. 1996. *Journal of Virology*. 70:8944-8960; Schiedner, G., et al. 1998. *Nature Genetics*. 18:180-183; Yang, Y., et al. 1994. *Proc. Natl. Acad. Sci. USA*. 91:4407-4411). Some of these proteins (e.g E4-orf4, pTP, or E3-11.6k) have pro-apoptotic activity (Langer, S. J., Schaak, J. 1996. *Virology*. 221:172-179; Lieber, A., et al. 1998. *J. Virology*. 72:9267-9277; Shtrichman, R., Kleinberger, T. 1998. *J. Virology*. 72:2975-2983; Tollefson, A. E., A et al. 1996 *J. Virology*. 70:2296-2306). Clearly, this would affect the outcome of transduction studies with Ad5GFP/F35, which allows for the efficient transfer of viral genomes into CD34+ cells implying significant expression of viral proteins. Moreover, recently published data indicate that short-term colony assay mostly measure mature progenitors and do not represent a rigorous test for transduction of potential stem cells.

A definitive demonstration that Ad5GFP/F35 based vectors can transduce HSC requires colony assays or preferably, repopulation assays in SCID-NOD mice. We can perform these studies with gutless vectors (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13) and integrating  $\square$ Ad.AAV vectors devoid of all viral genes (Lieber, A., et al. 1999. *J Virol* 73:9314-24) generated based on Ad5GFP/F35 chimeric capsids. Alternatively, gutless, retargeted vectors could be used to transiently express a retroviral receptor on CD34+ cells to increase their susceptibility to infection with retroviral vectors based on an approach that we have published earlier (Lieber, A., et al.. 1995. *Human Gene Therapy*. 6:5-11).

Our finding that Ad5GFP/F35 can efficiently transduce hematopoietic cells with potential stem cell capacity represents an important step towards stable gene transfer into HSCs and gene therapy of blood disorders. Furthermore, the virological aspects of this invention contribute to a better understanding of adenovirus cell interactions.

**I: Retargeting of Ad5 based vectors with modified fibers carrying specific ligand peptides for HSC and other cell types**

5 Another alternative to make Ad5-capsid-based vectors suitable for HSC gene therapy is to incorporate the coding sequence for HSC specific peptides into the H1 loop region of the Ad5 fiber gene. The modification of the H1-loop was successfully exercised by Krasnykh et al. with a 7 amino-acid long FLAG peptide (DYDDDDK). Using phage display peptide libraries (Pascualini, R. et al., 1996, *Nature*, 380:364-66), Renata  
10 Pasqualini (La Jolla Cancer Research Center) reported recently, at the First Meeting of the American Society for Gene Therapy, the identification of small peptide ligands specific for bone marrow cells. The corresponding sequences encoding these peptides can be added to modify the H1 loop sequence employing site-directed mutagenesis. Optimally, the ligands should allow for the efficient internalization of adenoviral particles  
15 based on a CAR- and integrin independent pathway. Modified adenoviral vectors containing the AAVBG cassette can be produced and tested for HSC tropism as described above.

**Adenovirus peptide display:**

20 In order to retarget adenoviruses to any cell type of interest, a strategy is provided which involves creating a library of adenoviruses displaying random peptides in their fiber knobs as ligands and screening this library for adenovirus variants with tropism to a particular cell type in vitro and potentially in vivo.

25 The development of the adenovirus peptide display technique is based on the following ideas. (i) Although the tertiary structure of the Ad5 fiber knob is known, it remains unclear which domains are involved in receptor binding. There are data suggesting that receptor-binding domains partially overlap with hemagglutination domains, which are  
30 well characterized for a number of serotypes. Therefore, three intramolecular loop regions representing potential receptor binding sites can be substituted by random peptide

libraries. Eight amino acid residues in the center of the FG, or GH loops can be substituted by octameric random peptides (Figs. 26 and 27). These substitutions will replace CAR tropism and allow for infection of refractory cell types. (ii) To synthesize the oligonucleotides encoding the peptide library a novel technique to assemble pre-synthesized trinucleotides representing the codons for all 20 amino acids is employed. This avoids termination codons and assures optimal codon usage and translation in human cells. Synthesis of a completely randomized library is possible with all 20 amino acids being incorporated with the same probability and a partially randomized library with only three (in average) random amino acids substitutions per octamer at random positions with a random amino acid to maintain certain critical features of the tertiary knob structure while introducing variability. The last model is based on the distribution of amino acids present in the hypervariable CDR 1 or 2 region of immunoglobulins. (iii) To maintain a representative library size of about  $10^{10}$  different octamers per modified loop, a new cloning strategy is employed to allow for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. This strategy is based on a "seamless" cloning technique available from Stratagen. (iv) In order to produce the library of viruses, viral genomic DNA containing the modified fiber sequences is transfected into 293 cells without reduction of the library size. This critical step is done by conjugating the viral library DNA to carrier Ad5-based adenovirus via poly-lysine to assure 100% transfection efficiency. This technique allows for coupling of  $\sim 1\mu\text{g}$  of plasmid DNA (or  $\sim 1 \times 10^{10}$  adenoviral genomes) to  $10^{10}$  viral particles which can be used to infect 293/cre cells at an MOI of 10-100. Importantly, the carrier adenoviral genome has the packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express cre recombinase (293/cre). This helper virus system is routinely used to produce so-called gutless adenoviruses. Therefore, the virus progeny represents library genomes packaged into capsids containing preferentially Ad5 fibers. This is important for the next infection step into 293 cells at a MOI of 1 to assure a homogeneous fiber population on the capsid where the fibers are encoded by the packaged genome.

30

#### J. Production Of Adenovirus Vectors With Increased Tropism To Hepatocytes



An example of a G-H loop substitution to target Ad5 to hepacytes was successful. Preliminary tests demonstrated that two evolutionarily conserved regions within the malaria circumsporozoite surface protein (CS) termed RI and RII+ mediate specific  
5 interaction with hepatocytes but not with other organs (including spleen, lung, heart and brain), nor with Kupffer cells, liver endothelial cells or with other regions of the hepatocyte membrane (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Shakibaei, M. and U. Frevert, 1996, *J. Exp. Med.*, 184:1699-711). These regions are conserved among different species including *Plasmodium berghei*, *P. cynomogli*, and *P. falciparum* that  
10 infect mouse, monkey and human hepatocytes, respectively (Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). Peptides derived from RI (KLKQPG) or RII (EWSPCSVTCGNGIQVRIK) blocked CS binding to hepatocytes and infection by sporozoites in vivo ((Cerami, C. et al., 1992, *Cell*, 70:1021-33; Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81). RI and RII+ peptides were  
15 separately inserted into Ad5-fiber knob (H-I and G-H loop) containing mutation with abolished binding to CAR and alpha-v integrins (Kirby, L. et al., 2000, *J. Virol.*, 74:2804-13; Wickham, T. J. et al., 1995, *Gene Ther.*, 2:750-6). Based on preliminary data, a short-shafted fiber was used so that the virus entry strategy predominantly depends on the interaction with the primary (hepatocyte-specific receptor). The hepatocyte-specific  
20 ligands are flanked by short glycine stretches to provide flexibility and embedded into a loop formed by two cystines. This is one of the classical strategies to incorporate ligands into a protein scaffold (Doi, N. and H. Yanagawa, 1998, *Cell Mol. Life Sci.*, 54:394-404; Koivunen, E. et al., 1995, *Biotechnology (NY)*, 13:265-70) and to guarantee their presentation at the protein surface. The biodistribution of the best variants is tested in  
25 vivo in C57B1/6 mice based on Southern blots or PCR for vector DNA in different organs. This mouse strain is known to be susceptible to infection with *P. berghei* (Chatterjee, S. et al., 1995, *Infect Immun.*, 63:4375-81).

### **K. Production Of Adenovirus Vectors With Increased Tropism To Tumor Cells**

A similar strategy is to insert two peptides obtained after selection for tumor tropism by displaying random peptides on filamentous phages. The first double cyclic peptide  
5 (RGD-4) proved to bind specifically to integrins present on tumor vasculature (Ellerby H. M. et al., 1999, *Nat. Med.*, 5:1032-8). The second peptide targets specific matrix metalloproteinases associated with metastatic tumor cells as shown for the breast cancer cell line MDA-MB-435 (Koivunen, E. et al., 1999, *Nat. Biotechnol.*, 17:768-74). Tropism-modified vectors are tested in animal models with hepatic metastases derived  
10 from MDA-MB-435 cells (Fig. 28).

### **L. Development Of A Peptide Display Technique Based On Adenoviruses**

A synthetic peptide library is described that allows adenovirus vectors to express random  
15 peptides in the G-H loop of the fiber knob domain. The technique of a phage display library is optimized to generate a library of adenoviruses displaying random peptides in their fiber knob. This library of adenovirus variants is then screened for tropism to a particular cell type in vitro and potentially in vivo. The oligonucleotides encoding the peptide library employ a novel technique to assemble pre-synthesized trinucleotides  
20 representing the codons for all 20 amino acids. This will end the termination codons and assure optimal codon usage and translation in human cells. To maintain a representative library size, a new "seamless" cloning strategy that allows for insertion of the library into the wild-type Ad5 genome without introducing additional amino acids at the substitution site and without transformation into bacteria. Transfection into 293 cells is done by  
25 conjugating the viral library DNA to carrier Ad5-based adenovirus via polylysine to assure a 100% transfection efficiency. Importantly, the carrier adenoviral genome has its packaging signal flanked by lox sites preventing the packaging of carrier viral DNA after infection of 293 cells that express Cre recombinase (293/cre). The library is produced with E1-positive viruses depleted for CAR and integrin tropism. Only variant that have  
30 successfully infected the cell type of interest will replicate, resulting in de novo produced

virus. The sequence of the peptide ligand that conferred the particular tropism will then be analyzed in do novo produced virus.

### EXAMPLE III

5

#### COMBINATION NOVEL ADENOVIRAL VECTOR AND MODIFIED FIBER PROTEIN

This example describes the following studies which combine the technology of the integrating adenovirus vector that is devoid of all adenoviral genes with the modified fiber protein that retargets the vector to quiescent HSC.

##### A. Transduction studies with re-targeted vectors in HSC:

15 In order to transduce quiescent HSC and integrate into chromosomal DNA, retargeted  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors are tested for reporter gene expression, and vector integration simultaneously while analyzing their clonogenic capacity. The modified  $\Delta\text{Ad.AAV}^{\text{fx}}$  hybrid vectors contain genomes devoid of all adenoviral genes (a "gutless" adenovirus vector) packaged into Ad5 capsids with modified fibers. Rep may be incorporated into  
20 these  $\Delta\text{Ad.AAV}^{\text{fx}}$  vectors to allow for site-specific integration into AAVS1.

##### Transduction studies:

Purified human CD34+ cells in IMDM/FCS+IL-3 and SCF are infected with different  
25 doses of  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-BG}$  ( $1\text{-}10^7$  genomes per cell). CD34+ cells infected with  $\Delta\text{Ad.AAV}^{\text{fx}}\text{-}\beta\text{Gal}$  are cultured for 2 days in suspension and sort  $\beta\text{-Gal}^+$  cells by FACS using FDG as substrate. This determines the infection efficiency.  $\beta\text{-gal}$  expressing cells are then submitted to clonogenic assays in semi-solid cultures (in two dishes per MOI) in the presence of multiple cytokines. (IL-3, SCF, Epo, G-CSF, GM-CSF, IL-7, Tpo). A  
30 first set of semi-solid cultures can be evaluated after 7 days; another set can be analyzed after 14 days. Colonies that have formed in semisolid culture can be characterized by

light microscopy and subsequently stained with X-Gal staining. Most of the vector genomes should remain episomal and can be lost with successive cell divisions. Thus, while most cells can be X-Gal positive at day 2 or day 7 after infection, most of the larger colonies (analyzed at day 14 p.i.) may not stain homogeneously for  $\beta$ -Gal. A representative number of X-Gal positive and X-Gal negative colonies can be picked and analyzed for episomal and integrated vector DNA. The outcome depends on the MOI used for infection and the integration status of the vector. These studies determine whether hybrid vectors can infect primitive progenitors.

10 Detailed characterization of hybrid vector integration:

CD34+ cells can be infected with  $\Delta$ Ad.AAV<sup>fx</sup>-SNori (MOI 1-10<sup>7</sup>) and subjected to G418 selection in methyl cellulose (MC) cultures in the presence of growth factors (IL-3 and SCF). The resulting colonies are a mixture of mainly myeloid cells. The number and morphology of G418 resistant colonies can be determined after 2 weeks of selection. This strategy may be disadvantageous in that the appropriate stem cell may not divide and form G418 resistant colonies under the specific culture condition used. Moreover, it may be difficult to perform G418 selection on a population of heterogeneous cells, which vary in their sensitivity to G418. Therefore, another set of  $\Delta$ Ad.AAV<sup>fx</sup>-SNori infected CD34+ cells can be cultured in methyl-cellulose (+IL-3, SCF) without G418 selection. After 2-3 weeks, single colonies can be picked from both (w/ and w/o G418) MC cultures, morphologically characterized, and analyzed for integrated vector using the modified protocol developed for integration studies in a small number of cells (see Figure 8). This strategy allows the assessment of whether hybrid vectors integrate into the genome of CD34+ cells cultured in the presence of growth factors. This study gives us an idea about potential position effects affecting neo or  $\beta$ gal expression from integrated vector copies and about the structure of the integrated vector and the flanking chromosomal regions.

30 An alternative method to confirm vector integration:

Fluorescence *in situ* hybridization (FISH) analysis, can be performed in individual cells from MC colonies. CD34+ cells are cultured in MC in the presence of growth factors to induce cell division and subsequently treated with colchicine. Metaphase chromosome spreads are analyzed with biotin-ATP labeled probe specific for the  $\beta$ Gal or SNori gene and a dioxigenin-UTP labeled probe for the human X-chromosome as an internal control (provided by Christine Disteché, University of Washington). Specific hybridization can be visualized with corresponding anti-biotin or anti-DIG antibodies labeled with different fluorochromes (e.g. FITC and Texas Red). Hybrid vector DNA may integrate as concatemers, which would facilitate detection by FISH. This technique allows one to localize the chromosomal integration sites of hybrid vectors.

Test transduction into quiescent bone marrow subpopulations:

Hybrid vectors described so far can be tested to see whether quiescent CD34+ cells can be stably transduced. To avoid significant cell proliferation, purified CD34+ cells are cultured in serum free IMDM supplemented with thrombopoietin (Tpo). Tpo can alone support the survival of stem cells without stimulating their active cell proliferation (Matsunaga, T. et al., 1998, *Blood*, 92:452-61; Papayannopoulou, T. et al., 1996, *Experimental Hematology*, 24:660-69). To analyze the proliferation status of CD34+ cells at the time point of infection with the hybrid vector  $\Delta$ Ad.AAV<sup>h</sup>BG, BrdU is added 2 hours before infection to the culture medium. One set of cells are maintained as suspension culture in IDAM containing Tpo only for two days. Another set of cells are grown in IDAM+Tpo supplemented with multiple cytokines. Forty eight hours after infection, CD34+ cells can be FACS sorted for beta Gal expression using FDG. FDG positive cells can be further analyzed for cellular DNA replication based on BrdU incorporation and for specific CD34+ subset markers. To do this, cytopins from FDG+ cells can be submitted to immunofluorescence with BrdU specific antibodies and with antibodies to specific cell surface markers (e.g. CD38, CD41). Alternatively, consecutive paraffin sections of the same cell can be analyzed for (a) transgene expression by X-Gal staining, (b) DNA synthesis based on BrdU incorporation, and (c) specific surface markers. This allows one to confirm that the culture conditions with Tpo alone prevent

significant genomic DNA replication and subsequent cell proliferation as well as to determine whether quiescent CD34+ cells can be infected based on beta Gal expression in cells where BrdU labeling is absent.

5 Test hybrid vectors integration into quiescent CD34+ cells:

Two sets of CD34+ cells are infected. The first set of  $\Delta$ Ad.AAV<sup>fx</sup>SNori infected cells are cultured for 5-7 days in the presence of cytokines; the other set is cultured without cytokines. To maintain CD34+ cell viability without cytokines during this period, the  
10 cells are cultured in the presence of Tpo or underlaid with a stromal cell line (AFT024) (Moore, K. et al., 1997, *Blood*, 89:4337-47), which can maintain HSC viable for 4 to 7 weeks. After this specific time period, both sets are submitted to clonogenic assays (in the presence of multiple cytokines) either in combination with G418 selection or without selection. Single colonies are analyzed morphologically and submitted to genomic DNA  
15 analysis (Figure 8) to determine the vector integration status. The ultimate proof for stem cell transduction is the in vivo survival/expansion assay. To do this, the CD34+ cells expressing beta Gal are used for transplantation experiments. If the number of FDG+ cells is not sufficient, total  $\Delta$ Ad.AAV<sup>fx</sup>BG infected cells as well as all  $\Delta$ Ad.AAV<sup>fx</sup>-SNori infected cells can be used directly without selection. Transplantation can be performed  
20 via tail vein injection into sublethally irradiated SCID NOD mice (Dao, M.A., et al., 1998, *Blood*, 4, 1243-1255; Matsunaga, T. et al., 1998, *Blood*, 92:452-61). At different time points after transplantation (4 to 8 weeks), mice can be sacrificed to obtain bone marrow cells which then can be cultured in suspension until various assays are performed for X-Gal and cell markers as described earlier. These cells also can be submitted to a  
25 secondary colony assay in MC or secondary transplantation into SCID NOD mice. Furthermore, MC colonies derived from these cells can be analyzed for the presence of integrated vector DNA by the method illustrated in Figure 8. The expression and integration data together allow conclusions about the repopulation efficiency and about potential position effects.

30

**B. Optimization of  $\Delta$ Ad.AAV<sup>fx</sup> vectors for  $\gamma$  r-globin expression in hematopoietic stem cells:**

One specific example of the invention is (a) to construct retargeted hybrid vector with the  $\gamma$ -globin as the transgene under the control of erythroid cell specific promoter, (b) to analyze the level and kinetics of  $\gamma$ -globin expression after transduction with hybrid vectors in *in vitro* and *in vivo* assays, (c) if required, to protect gene expression from position effects using  $\gamma$ -globin LCRs or insulators incorporated into hybrid vectors, and (d) to study whether  $\gamma$ -globin introns or heterologous introns can increase  $\gamma$ -globin expression.

Another central issue of the invention is to demonstrate that hybrid vectors can accommodate larger transgenes than rAAV and retroviruses. The insert size limitation of these vectors is 5kb. Transgene cassettes up to 8kb can be inserted into hybrid vectors as described. The maximal insert size may be about 14kb, if hybrid vectors are produced on the basis of E2a and/or E4 deleted rAd vectors in corresponding packaging cell lines. The maximal insert size in hybrid vectors is dictated by the packaging limit of first generation vectors (Ad.AAV) (<36kb) which are necessary intermediates for hybrid virus production at large scale. It is expected that stability and titer of Ad.AAV vectors with an 8kb globin gene cassette is comparable to the vector containing the 2.5-3.5kb cassette used in Ad.AAVBG, Ad.AAV1, and Ad.AAVSNori. The following example experiments address these issues.

**Production of  $\Delta$ Ad.AAV<sup>fx</sup> with large globin expression cassettes:**

In order to improve the condition of sickle cell disease, the expression level of the transferred  $\gamma$ -lobin gene must be at least 50% of that of each endogenous  $\beta$ gene. These levels of transgene expression can only be achieved by using optimal expression cassettes, including extended LCRs and intron containing gamma genes. (Forrester, W.C., et al., 1986, *Proc. Natl. Acad. Sci. USA* 83, 1359-1363; Fraser, P., et al., 1998, *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in*

*Hematology*, 35, 105-111; Martin, D. et al., 1996, *Current Opinion in Genetics and Development*, 6:488-95), So far, most of the  $\gamma$ -globin expression cassettes are designed for retroviral and rAAV vectors, thus, less than 5kb and have to be devoid of internal splice sites or poly adenylation signals. With integrating vectors described herein, it is possible to go beyond this size limitation. This allows one to improve  $\gamma$ -globin expression in bone marrow cells in terms of an adequate expression level and long term persistence. For this purpose,  $\gamma$ -globin constructs developed by Li et al (Emery, D. W., et al. 1999 *Hum Gene Ther* 10:877-88; Li, Q., et al. 1999. *Blood* 93:2208-16) or by Ellis et al (Ellis, J., et al., 1996, *EMBO J.*, 15, 562-568; Ellis, J., et al., 1997, *Nucleic Acids Res.* 25, 1296-1302) is chosen.

(i) The first cassette contains a  $\gamma$ -globin expression unit used in retroviral vectors. This allows for a direct comparison between the two systems. This construct includes the beta promoter from -127 to the beta initiation codon, which is connected in frame with the gamma coding region. This beta promoter is combined with the 300bp HS40 derived from the human alpha globin locus, which acts as a strong enhancer for globin expression. The globin gene is the 1.1kb version with intron 1 and partially deleted intron 2. A second cassette is generated containing the HS40 beta promoter and gamma globin gene with the complete 3.3kb gamma globin gene.

(ii) The second construct contains the 6.5kb beta  $\mu$ LCR, which confer a dominant chromatin opening activity and an adequate level of gamma globin expression in transgenic mice. The LCR is linked to the short 1.1kb version of the gamma globin gene or the complete 3.3kb gamma gene.

(iii) Additional globin expression cassette can be generated which include insulators, MARs or SARs, as well as other elements that can improve transgene expression from integrated vectors or in transgenic animals, like introns derived from the HPRT or hGH genes (Chung, J.H., et al., 1997, *Proc. Natl. Acad. Sci. USA* 94, 575-580; Dunaway, M., et al, 1993, *Mol. Cell. Biol.*, 17, 182-189; Felsenfeld, G., et al., 1996,



*Proc. Natl. Acad. Sci. USA* 93, 93840-9388; Klehr, D., et al., 1991, *Biochemistry*, 30, 1264-1270).

Transduction studies with  $\Delta$ Ad.AAVfx-globin vectors:

5

Transduction studies with globin-hybrid vectors are performed as described earlier (Steinwaerder, D. S., et al. 1999. *J Virol* 73:9303-13). Transduced CD34+ cells are submitted to differentiation in colony assays or analyzed *in vivo* expansion assays in SCID-NOD mice. MC-colonies or bone marrow cells from experimental mice are analyzed for globin expression. Gamma-globin expression is measured using fluorescent anti-gamma-globin antibodies. RNAase protection studies can be performed to specifically quantitate gamma globin mRNA in comparison with  $\gamma$ -globin RNA. For these studies around  $10^4$ - $10^5$  cells are needed per test.

15 

Position effects:

In the absence of the LCR, globin genes are subjected to strong position effects when they are transferred into cultured CD34+ cells or erythroleukemic lines (Fraser, P., et al., 1998, *Curr. Opinion in Cell Bio.*, 10, 361-365; Grosveld, F., et al., 1998, *Seminars in Hematology*, 35, 105-111). Another concern is that site-specific integration of  $\Delta$ Ad.AAV/rep vectors into AAVS1 may silence transgene expression. If silencing happens, it can be overcome by incorporating LCRs such as the 6.5kb  $\gamma$  globin  $\mu$ LCR (Ellis, J., et al., 1996, *EMBO J.*, 15, 562-568; Grosveld, F., et al., 1998, *Seminars in Hematology*, 35, 105-111) or insulators into  $\Delta$ Ad.AAV based expression units. Insulators are DNA elements that protect an integrated reporter gene from chromosomal position effects or that block enhancer activated transcription from a downward promoter. Insulator elements are known for *Drosophila melanogaster* genes (Gypsy, suppressor of Hairy wing, scs, scs', Fab-7). for the chicken beta-globin gene (HS4) and for the T cell receptor (BEAD1; 14, 21.25). Specifically, the *Drosophila* gypsy or the beta globin insulator can be inserted as two copies flanking the globin expression cassette into hybrid vectors. The position effects can be examined in transduced MC-colonies based on the

30

analysis of integrated vector DNA (see Fig. 29) and gamma-globin mRNA quantification. Analogous studies can be performed on transduced human bone marrow cells obtained after transplantation of infected CD34+ cells into SCID-NOD mice.

5 Intron effects on gamma-globin expression:

A number of reports reveal that the deletion of globin introns, particularly the second intron of the beta and gamma genes, decrease globin mRNA stability and thus the expression level (Antoniou, M. et al., 1998, *Nucleic Acid Res.*, 26:721-9). RNA viruses  
10 such as onco-retro, lenti- and foami viruses are problematic as vehicles for intron-containing transgenes. Because  $\Delta$ Ad.AAV is a DNA virus, it should package globin introns and LCRs if necessary without the decreased titers and rearrangements observed with retroviral vectors.

## APPENDIX I

### HUMAN AND ANIMAL ADENOVIRUSES AVAILABLE FROM AMERICAN TYPE CULTURE COLLECTION

5

- 1: Adenovirus Type 21 ATCC VR-1099 (NIAID V-221-002-014)
- 2: SA18 (Simian adenovirus 18) ATCC VR-943 Classification
- 10 3: SA17 (Simian adenovirus 17) ATCC VR-942 Classification
- 4: Adenovirus Type 47 ATCC VR-1309 Classification: Adenov
- 5: Adenovirus Type 44 ATCC VR-1306 Classification: Adenov
- 6: Avian adenovirus Type 4 ATCC VR-829 Classification: Ad
- 7: Avian adenovirus Type 5 ATCC VR-830 Classification: Ad
- 15 8: Avian adenovirus Type 7 ATCC VR-832 Classification: Ad
- 9: Avian adenovirus Type 8 ATCC VR-833 Classification: Ad
- 10: Avian adenovirus Type 9 ATCC VR-834 Classification: Ad
- 11: Avian adenovirus Type 10 ATCC VR-835 Classification: A
- 12: Avian adenovirus Type 2 ATCC VR-827 Classification: Ad
- 20 13: Adenovirus Type 45 ATCC VR-1307 Classification: Adenov
- 14: Adenovirus Type 38 ATCC VR-988 Permit: PHS permit requ
- 15: Adenovirus Type 46 ATCC VR-1308 Classification: Adenov
- 16: Simian adenovirus ATCC VR-541 Classification: Adenovir
- 17: SA7 (Simian adenovirus 16) ATCC VR-941 Classification:
- 25 18: Frog adenovirus (FAV-1) ATCC VR-896 Classification: Ad
- 19: Adenovirus Type 48 (candidate) ATCC VR-1406 Classifica
- 20: Adenovirus Type 42 ATCC VR-1304 Classification: Adenov

- 21: Adenovirus type 49 (candidate) ATCC VR-1407 Classification:
- 22: Adenovirus Type 43 ATCC VR-1305 Classification: Adenovirus
- 23: Avian adenovirus Type 6 ATCC VR-831 Permit: USDA permit
- 24: Avian adenovirus Type 3 (Inclusion body hepatitis virus)
- 5 25: Bovine adenovirus Type 3 ATCC VR-639 Classification: A
- 26: Bovine adenovirus Type 6 ATCC VR-642 Permit: USDA permit
- 27: Canine adenovirus ATCC VR-800 Classification: Adenovirus
- 28: Bovine adenovirus Type 5 ATCC VR-641 Permit: USDA permit
- 29: Adenovirus Type 36 ATCC VR-913 Classification: Adenovirus
- 10 30: Ovine adenovirus type 5 ATCC VR-1343 Classification: A
- 31: Adenovirus Type 29 ATCC VR-272 Classification: Adenovirus
- 32: Swine adenovirus ATCC VR-359 Classification: Adenovirus
- 33: Bovine adenovirus Type 4 ATCC VR-640 Permit: USDA permit
- 34: Bovine adenovirus Type 8 ATCC VR-769 Permit: USDA permit
- 15 35: Bovine adenovirus Type 7 ATCC VR-768 Permit: USDA permit
- 36: Adeno-associated virus Type 2 (AAV-2H) ATCC VR-680 Classification:
- 37: Adenovirus Type 4 ATCC VR-4 Classification: Adenovirus
- 38: Adeno-associated virus Type 3 (AAV-3H) ATCC VR-681 Classification:
- 39: Peromyscus adenovirus ATCC VR-528 Classification: Adenovirus
- 20 40: Adenovirus Type 15 ATCC VR-661 Classification: Adenovirus
- 41: Adenovirus Type 20 ATCC VR-662 Classification: Adenovirus
- 42: Chimpanzee adenovirus ATCC VR-593 Classification: Adenovirus
- 43: Adenovirus Type 31 ATCC VR-357 Classification: Adenovirus
- 44: Adenovirus Type 25 ATCC VR-223 Classification: Adenovirus

25

- 45: Chimpanzee adenovirus ATCC VR-592 Classification: Aden  
46: Chimpanzee adenovirus ATCC VR-591 Classification: Aden  
47: Adenovirus Type 26 ATCC VR-224 Classification: Adenovi  
48: Adenovirus Type 19 ATCC VR-254 Classification: Adenovi  
5 49: Adenovirus Type 23 ATCC VR-258 Classification: Adenovi  
50: Adenovirus Type 28 ATCC VR-226 Classification: Adenovi  
51: Adenovirus Type 6 ATCC VR-6 Classification: Adenovirus  
52: Adenovirus Type 2 Antiserum: ATCC VR-1079 AS/Rab (NIA  
53: Adenovirus Type 6 ATCC VR-1083 (NIAID V-206-003-014)  
10 54: Ovine adenovirus type 6 ATCC VR-1340 Classification: A  
55: Adenovirus Type 3 ATCC VR-847 (Derived from NIAID V-20)  
56: Adenovirus Type 7 ATCC VR-7 Classification: Adenovirus  
57: Adenovirus Type 39 ATCC VR-932 Classification: Adenovi  
58: Adenovirus Type 3 ATCC VR-3 Classification: Adenovirus  
15 59: Bovine adenovirus Type 1 ATCC VR-313 Classification: A  
60: Adenovirus Type 14 ATCC VR-15 Classification: Adenovir  
61: Adenovirus Type 1 ATCC VR-1078 (NIAID V-201-001-014)  
62: Adenovirus Type 21 ATCC VR-256 Classification: Adenovi  
63: Adenovirus Type 18 ATCC VR-1095 (NIAID V-218-003-014)  
20 64: Baboon adenovirus ATCC VR-275 Classification: Adenovir  
65: Adenovirus Type 10 ATCC VR-11 Classification: Adenovir  
66: Adenovirus Type 33 ATCC VR-626 Classification: Adenovi

- 67: Adenovirus Type 34 ATCC VR-716 Classification: Adenovi  
68: Adenovirus Type 15 ATCC VR-16 Classification: Adenovir  
69: Adenovirus Type 22 ATCC VR-257 Classification: Adenovi  
70: Adenovirus Type 24 ATCC VR-259 Classification: Adenovi  
5 71: Adenovirus Type 17 ATCC VR-1094 (NIAID V-217-003-014)  
72: Adenovirus Type 4 ATCC VR-1081 (NIAID V-204-003-014)  
73: Adenovirus Type 16 ATCC VR-17 Classification: Adenovir  
74: Adenovirus Type 17 ATCC VR-18 Classification: Adenovir  
75: Adenovirus Type 16 ATCC VR-1093 (NIAID V-216-003-014)  
10 76: Bovine adenovirus Type 2 ATCC VR-314 Classification: A  
77: SV-30 ATCC VR-203 Classification: Adenovirus, Simian (  
78: Adenovirus Type 32 ATCC VR-625 Classification: Adenovi  
79: Adenovirus Type 20 ATCC VR-255 Classification: Adenovi  
80: Adenovirus Type 13 ATCC VR-14 Classification: Adenovir  
15 81: Adenovirus Type 14 ATCC VR-1091 (NIAID V-214-001-014)  
82: Adenovirus Type 18 ATCC VR-19 Classification: Adenovir  
83: SV-39 ATCC VR-353 Classification: Adenovirus, Simian (  
84: Adenovirus Type 11 ATCC VR-849 (Derived from NIAID V-2  
85: Duck adenovirus (Egg drop syndrome) ATCC VR-921 Permi  
20 86: Adenovirus Type 1 ATCC VR-1 Classification: Adenovirus  
87: Chimpanzee adenovirus ATCC VR-594 Classification: Aden  
88: Adenovirus Type 15 ATCC VR-1092 (NIAID V-215-003-014)

- 89: Adenovirus Type 13 ATCC VR-1090 (NIAID V-213-003-014)  
90: Adenovirus Type 8 ATCC VR-1368 (Derived from NIAID V-20  
91: SV-31 ATCC VR-204 Classification: Adenovirus, Simian (  
92: Adenovirus Type 9 ATCC VR-1086 (NIAID V-209-003-014)  
5 93: Mouse adenovirus ATCC VR-550 Classification: Adenoviru  
94: Adenovirus Type 9 ATCC VR-10 Classification: Adenoviru  
95: Adenovirus Type 41 ATCC VR-930 Classification: Adenovi  
96: CI ATCC VR-20 Classification: Adenovirus, Simian (Mast  
97: Adenovirus Type 40 ATCC VR-931 Classification: Adenovi  
10 98: Adenovirus Type 37 ATCC VR-929 Classification: Adenovi  
99: Marble spleen disease virus (Hemorrhagic enteritis virus  
100: Adenovirus Type 35 ATCC VR-718 Classification: Adenovi  
101: SV-32 (M3) ATCC VR-205 Classification: Adenovirus, Sim  
102: Adenovirus Type 28 ATCC VR-1106 (NIAID V-228-003-014)  
15 103: Adenovirus Type 10 ATCC VR-1087 (NIAID V-210-003-014)  
104: Adenovirus Type 20 ATCC VR-1097 (NIAID V-220-003-014)  
105: Adenovirus Type 21 ATCC VR-1098 (NIAID V-221-011-014)  
106: Adenovirus Type 25 ATCC VR-1103 (NIAID V-225-003-014)  
107: Adenovirus Type 26 ATCC VR-1104 (NIAID V-226-003-014)  
20 108: Adenovirus Type 31 ATCC VR-1109 (NIAID V-231-001-014)  
109: Adenovirus Type 19 ATCC VR-1096 (NIAID V-219-002-014)  
110: SV-36 ATCC VR-208 Classification: Adenovirus, Simian (

- 111: SV-38 ATCC VR-355 Classification: Adenovirus, Simian (  
112: SV-25 (M8) ATCC VR-201 Classification: Adenovirus, Sim  
113: SV-15 (M4) ATCC VR-197 Classification: Adenovirus, Sim  
114: Adenovirus Type 22 ATCC VR-1100 (NIAID V-222-003-014)  
5 115: SV-23 (M2) ATCC VR-200 Classification: Adenovirus, Sim  
116: Adenovirus Type 11 ATCC VR-12 Classification: Adenovir  
117: Adenovirus Type 24 ATCC VR-1102 (NIAID V-224-003-014)  
118: Avian adenovirus Type 1 (Chicken Embryo Lethal Orphan: C  
119: SV-11 (M5) ATCC VR-196 Classification: Adenovirus, Sim  
10 120: Adenovirus Type 5 ATCC VR-5 Classification: Adenovirus  
121: Adenovirus Type 23 ATCC VR-1101 (NIAID V-223-003-014)  
122: SV-27 (M9) ATCC VR-202 Classification: Adenovirus, Sim  
123: Avian adenovirus Type 2 (GAL) ATCC V-280 Classificati  
124: SV-1 (M1) ATCC VR-195 Classification: Adenovirus, Simi  
15 125: SV-17 (M6) ATCC VR-198 Classification: Adenovirus, Sim  
126: Adenovirus Type 29 ATCC VR-1107 (NIAID V-229-003-014)  
127: Adenovirus Type 2 ATCC VR-846 Classification: Adenovir  
128: SV-34 ATCC VR-207 Classification: Adenovirus, Simian (  
129: SV-20 (M7) ATCC VR-199 Classification: Adenovirus, Sim  
20 130: SV-37 ATCC VR-209 Classification: Adenovirus, Simian (  
131: SV-33 (M10) ATCC VR-206 Classification: Adenovirus, Si



- 132: Avian adeno-associated virus ATCC VR-865 Classificatio  
133: Adeno-associated (satellite) virus Type 4 ATCC VR-646  
134: Adenovirus Type 30 ATCC VR-273 Classification: Adenovi  
135: Adeno-associated (satellite) virus Type 1 ATCC VR-645  
5 136: Infectious canine hepatitis (Rubarth's disease, Fox ence  
137: Adenovirus Type 27 ATCC VR-1105 (NIAID V-227-003-014)  
138: Adenovirus Type 12 ATCC VR-863 (Derived from NIAID V-2  
139: Adeno-associated virus Type 2 (molecularly cloned) ATCC  
140: Adenovirus Type 7a ATCC VR-848 (Derived from NIAID V-2

10

What is claimed:

1. A first generation recombinant adenovirus vector a portion of which integrates into a host cell genome, comprising:
  - 5 a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
  - d) At least one adenoviral sequence which directs adenoviral replication; and
  - 10 e) A right adenoviral inverted terminal repeat sequence,wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.
- 15 2. The adenovirus vector of claim 1, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype.
3. The adenovirus vector of claim 1, wherein the sequence which directs adenoviral replication comprises a sequence on the anti-parallel strand which encodes an adenoviral fiber protein including a fiber tail, a fiber shaft, and a fiber knob,  
20 wherein the fiber knob includes a G-H loop region.
4. The adenovirus vector of claim 3, wherein the sequence on the anti-sense strand which encodes the fiber tail is from the same serotype as the adenoviral inverted  
25 repeat sequence.
5. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:
  - 30 a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;

- c) A transgene sequence 3' to the polyadenylation sequence;
- d) A promoter sequence 3' to the polyadenylation sequence; and
- e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.

5

6. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises:

- a) A left cassette inverted terminal repeat sequence;
- b) A promoter sequence 3' to the left inverted terminal repeat sequence;
- 10 c) A transgene sequence 3' to the promoter sequence;
- d) A polyadenylation sequence 3' to the transgene sequence; and
- e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.

15 7. The adenovirus vector of claim 5 or 6, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.

20 8. The adenovirus vector of claim 5 or 6, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.

25 9. The adenovirus vector of claim 8, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.

10. The adenovirus vector of claim 8, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.

30 11. The adenovirus vector of claim 8, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.

12. The transgene cassette of claim 5 or 6 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
- 5 13. The transgene cassette of claim 5 or 6, further comprising an insulator element.
14. The transgene cassette of claim 5 or 6, further comprising a bacterial origin of replication.
- 10 15. The adenoviral vector of claim 1, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
- 15 16. A first generation recombinant adenovirus vector which targets a host cell of interest and a portion of which integrates into the host cell genome so targeted, comprising two DNA strands, each strand being antiparallel to the other, the first strand comprising:
- 20 a) A left adenovirus inverted terminal repeat sequence;
- b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
- c) A transgene cassette sequence 3' to the adenoviral packaging sequence;
- d) At least one adenoviral sequence which directs adenoviral replication; and
- e) A right adenoviral inverted terminal repeat sequence,
- 25 wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome, and wherein the second strand comprises a sequence which encodes an adenoviral fiber protein that permits targeting of the vector into the host cell of interest.

17. The adenovirus vector of claim 16, wherein the adenoviral protein includes a fiber tail, a fiber shaft, and a fiber knob, wherein the fiber knob includes a G-H loop region.
- 5 18. The adenovirus vector of claim 16, wherein the left and right adenovirus inverted terminal repeat sequences and the packaging sequence are from the same adenoviral serotype.
- 10 19. The adenovirus vector of claim 17, wherein the fiber tail is from the same serotype as the left and right adenoviral inverted repeat sequences.
20. The adenovirus vector of claim 17, wherein the fiber shaft is from a different serotype as the left and right adenoviral inverted repeat sequences.
- 15 21. The adenovirus vector of claim 20, wherein the fiber shaft is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
22. The adenovirus vector of claim 17, wherein the fiber shaft comprises a shortened length.
- 20 23. The adenovirus vector of claim 17, wherein the fiber knob is from a different serotype as the left and right adenoviral inverted repeat sequences.
24. The adenovirus vector of claim 23, wherein the fiber knob is from a serotype selected from a group consisting of serotype 3, 7, 9, 11, and 35.
- 25 25. The adenovirus vector of claim 17, wherein the fiber knob is a modified fiber knob protein comprising the G-H loop replaced with a heterologous peptide ligand sequence which binds to at least one surface protein on the host cell of interest.
- 30

26. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 5 b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - 10 e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
27. The adenovirus vector of claim 16, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - 15 b) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - c) A transgene sequence 3' to the promoter sequence;
  - d) A polyadenylation sequence 3' to the transgene sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
  - 20
28. The adenovirus vector of claim 26 or 27, wherein the left and the right cassette inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 25 29. The adenovirus vector of claim 26 or 27, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.
- 30 30. The adenovirus vector of claim 24, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.

31. The adenovirus vector of claim 24, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentamycin.
- 5 32. The adenovirus vector of claim 24, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
- 10 33. The transgene cassette of claim 26 or 27 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
34. The transgene cassette of claim 26 or 27, further comprising an insulator element.
- 15 35. The transgene cassette of claim 26 or 27, further comprising a bacterial origin of replication.
- 20 36. The adenoviral vector of claim 16, wherein the adenoviral sequences which direct adenoviral replication are selected from a group consisting of E2 and E4; E1, E2 and E4; E2 and E4; and E2, E3, and E4.
- 25 37. A recombinant gutless adenovirus vector a portion of which integrates into a host cell genome, comprising:
- a) A left adenovirus inverted terminal repeat sequence;
  - b) An adenoviral packaging sequence 3' to the left adenovirus inverted terminal repeat sequence;
  - c) A transgene cassette sequence 3' to the adenoviral packaging sequence; and
  - d) A right adenoviral inverted terminal repeat sequence,
- 30 wherein the left and right terminal repeat sequences permit integration of the transgene cassette sequence into the host cell genome.

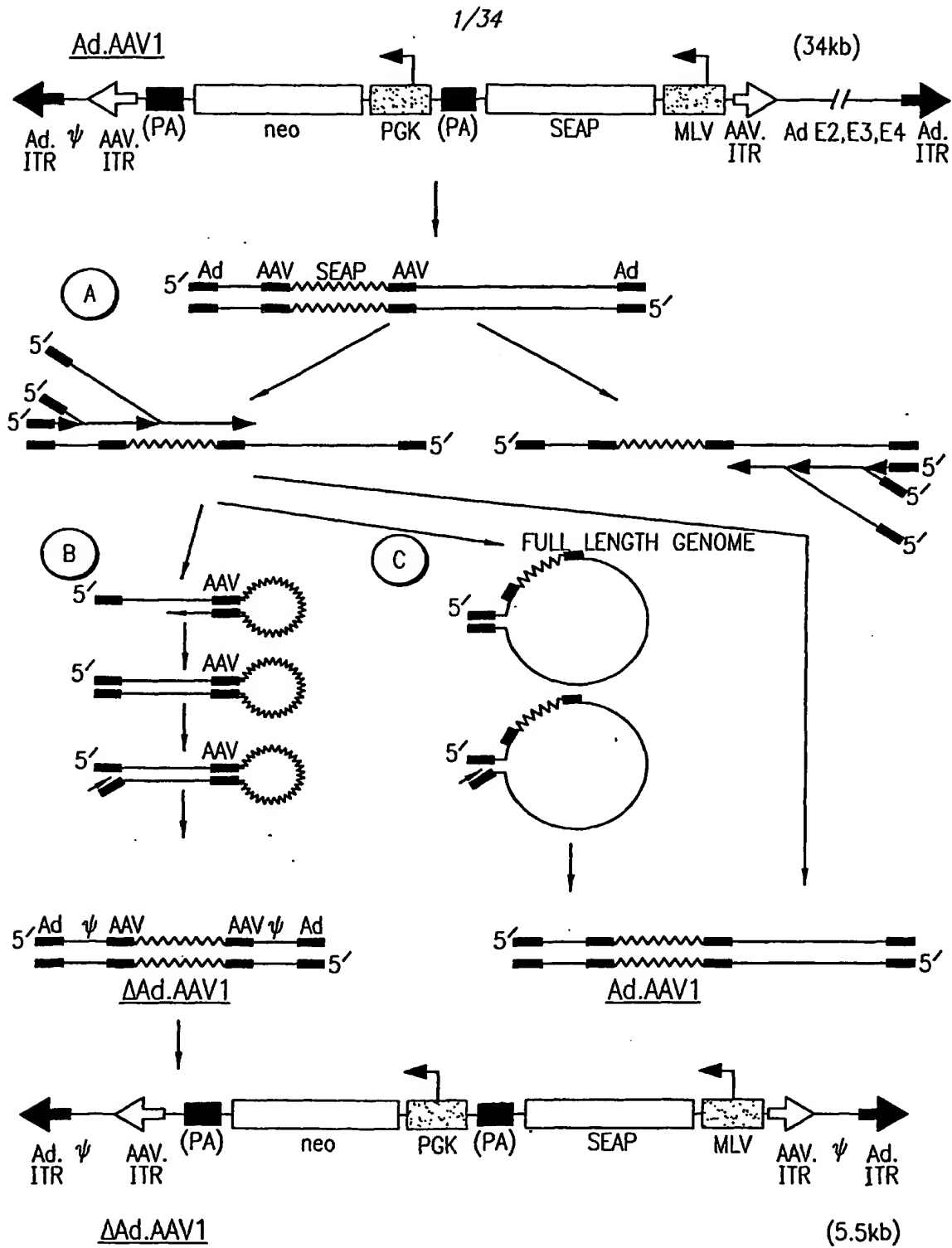
38. The adenovirus vector of claim 36, wherein the left and right adenovirus inverted repeat sequence and the packaging sequence are from the same adenoviral serotype. *(this is the base vector)*
- 5 39. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- a) A left cassette inverted terminal repeat sequence;
  - b) A polyadenylation sequence 3' to the left inverted terminal repeat sequence;
  - 10 c) A transgene sequence 3' to the polyadenylation sequence;
  - d) A promoter sequence 3' to the polyadenylation sequence; and
  - e) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 15 40. The adenovirus vector of claim 36, wherein the transgene cassette sequence comprises:
- b) A left cassette inverted terminal repeat sequence;
  - c) A promoter sequence 3' to the left inverted terminal repeat sequence;
  - d) A transgene sequence 3' to the promoter sequence;
  - 20 e) A polyadenylation sequence 3' to the transgene sequence; and
  - f) A right cassette inverted terminal repeat sequence 3' to the polyadenylation sequence.
- 25 41. The adenovirus vector of claim 38 or 39, wherein the left and the right inverted terminal repeat sequences each comprise an adenoviral-associated inverted terminal repeat sequence.
- 30 42. The adenovirus vector of claim 38 or 39, wherein the transgene sequence is selected from a group consisting of a therapeutic gene, a selectable gene, and a reporter gene.



43. The adenovirus vector of claim 41, wherein the therapeutic gene is selected from a group consisting of gamma globin, and human alpha-1 anti-trypsin.
44. The adenovirus vector of claim 41, wherein the selectable gene is selected from a group consisting of neomycin, ampicillin, penicillin, tetracycline, and gentomycin.
45. The adenovirus vector of claim 41, wherein the reporter gene is selected from a group consisting of green fluorescent protein, beta galactosidase, alkaline phosphatase.
46. The adenovirus vector of claim 38 or 39 further comprising an inverted repeat sequence located 3' to the left inverted terminal repeat sequence or located 5' to the right inverted terminal repeat sequence.
47. The adenovirus vector of claim 38 or 39, further comprising an insulator element.
48. The adenovirus vector of claim 38 or 39, further comprising a bacterial origin of replication.
49. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 5' portion of a gene of interest.
50. The adenovirus vector of claim 1, wherein the transgene cassette sequence comprises a 3' portion of a gene of interest.
51. A method of producing a resolved gutless adenovirus vector in a suitable cell, said method comprising introducing a first and a second adenovirus vectors of claim 1 or 16 into the cell under suitable conditions so that the recombinant adenovirus vectors undergo homologous recombination thereby producing a resolved gutless adenovirus vector.

52. A resolved gutless adenovirus vector produced by the method of claim 50.
53. The method of claim 50, wherein the first adenovirus vector comprises a transgene cassette having a 5' portion of a gene of interest, and wherein the  
5 second adenovirus vector comprises a transgene cassette having a 3' portion of the gene of interest, and wherein a part of the 5' portion overlaps with a part of the 3' portion so that homologous recombination occurs.
54. A method of producing a resolved gutless recombinant Ad vector by homologous  
10 recombination in a suitable cell, said method comprising contacting two parental recombinant Ad vectors, each comprising a transgene cassette containing a portion of a selected transgene with a region of overlapping homology, so that the first and second parental recombinant Ad vectors undergo homologous recombination at the region of overlapping homology, resulting in a resolved  
15 recombinant gutless Ad vector having both portions of the selected transgene, and wherein the selected transgene is within a transgene cassette flanked by a pair of ITRs.
55. A resolved gutless adenovirus vector produced by the method of claim 53.
- 20 56. An adenovirus library comprising a plurality of adenovirus vectors expressing fiber proteins which are displayed and modified by random peptide insertions.
57. The library of claim 55, wherein said fiber protein so displayed comprises a  
25 random peptide substituted in the G-H loop of the fiber protein knob domain.
58. A screening method for targeting adenovirus vectors for gene therapy comprising  
30 contacting the adenovirus library of claim 55 with a plurality of cells so that the cells are transduced with the adenovirus vectors of the adenovirus library transduction occurs and detecting the cells so transduced.

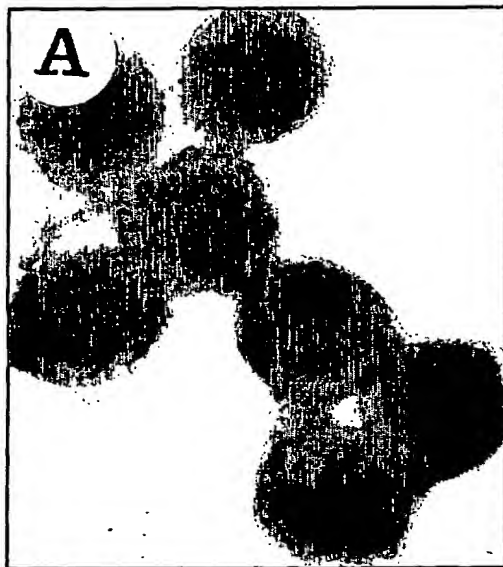
59. The adenovirus vector of claim 5, 6, 26, 27, 38, or 39 further comprising a nucleotide sequence encoding a rep78 protein.



**FIG. 1**  
SUBSTITUTE SHEET (RULE 26)

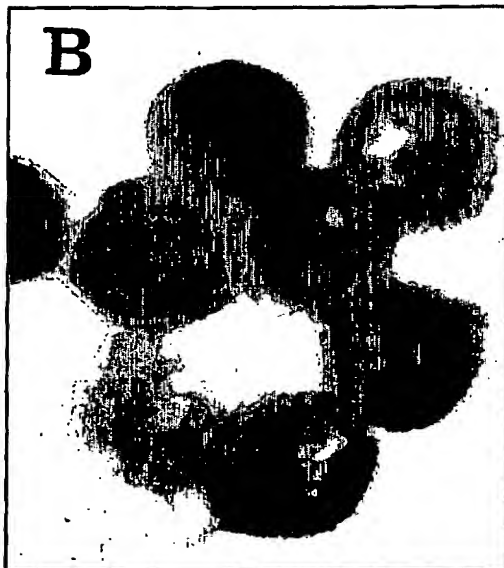
2/34

Ad.AAV  
(33kb)



50nm

$\Delta$ Ad.AAV  
(5.5kb)



50nm

*FIG. 2*

SUBSTITUTE SHEET (RULE 26)

3/34

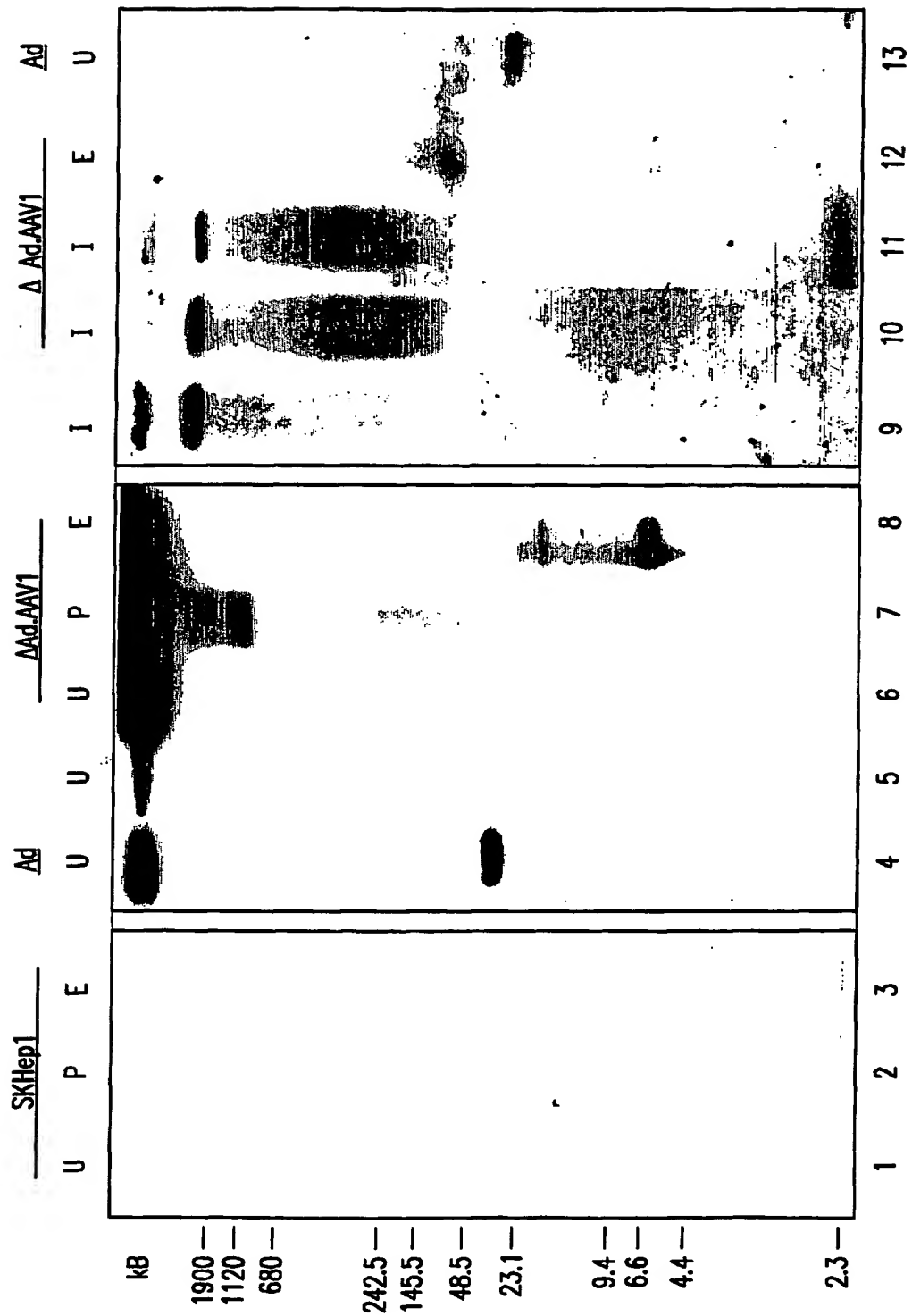


FIG. 3

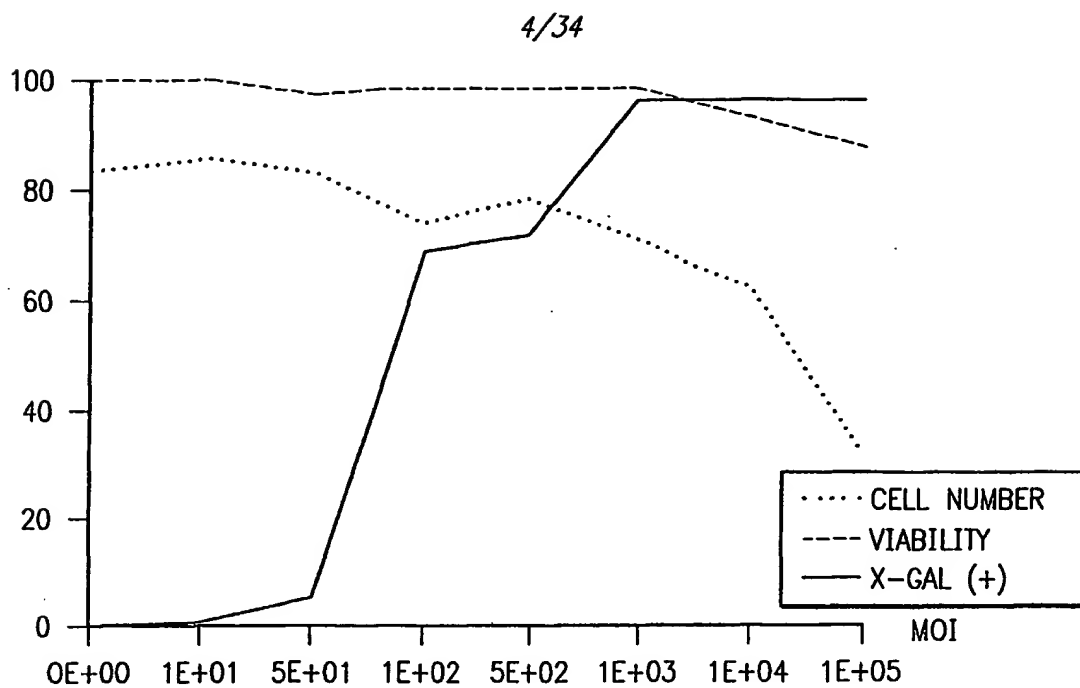


FIG. 4A K562

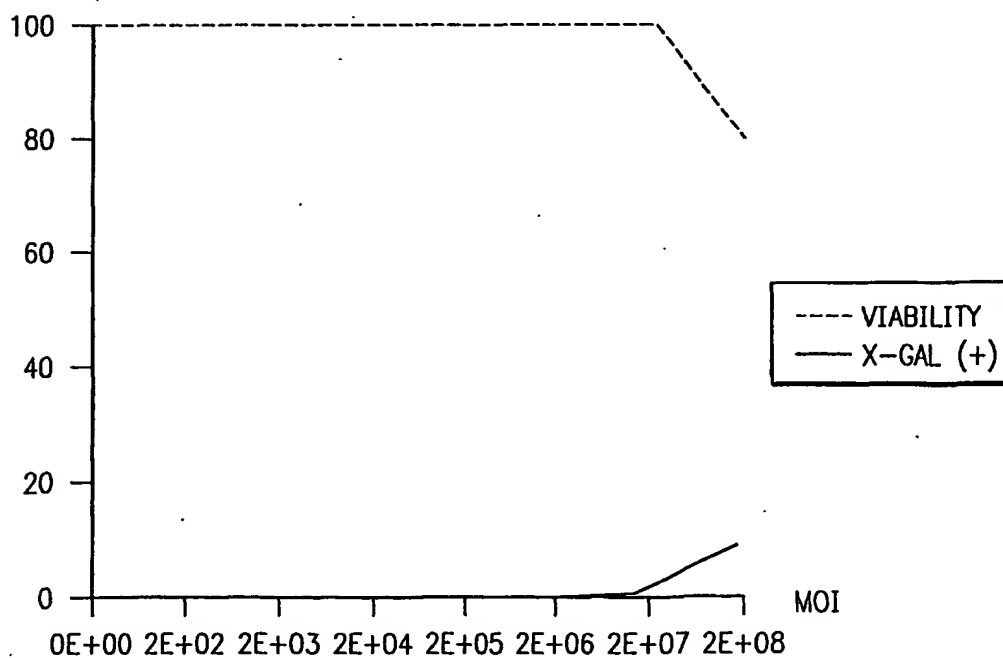


FIG. 4B CD34+

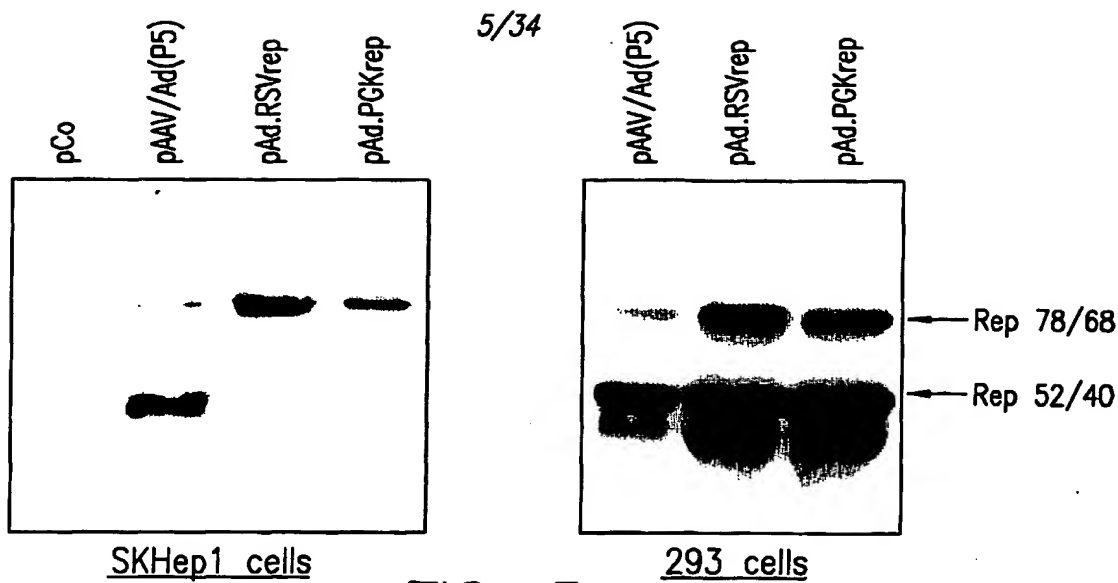


FIG. 5

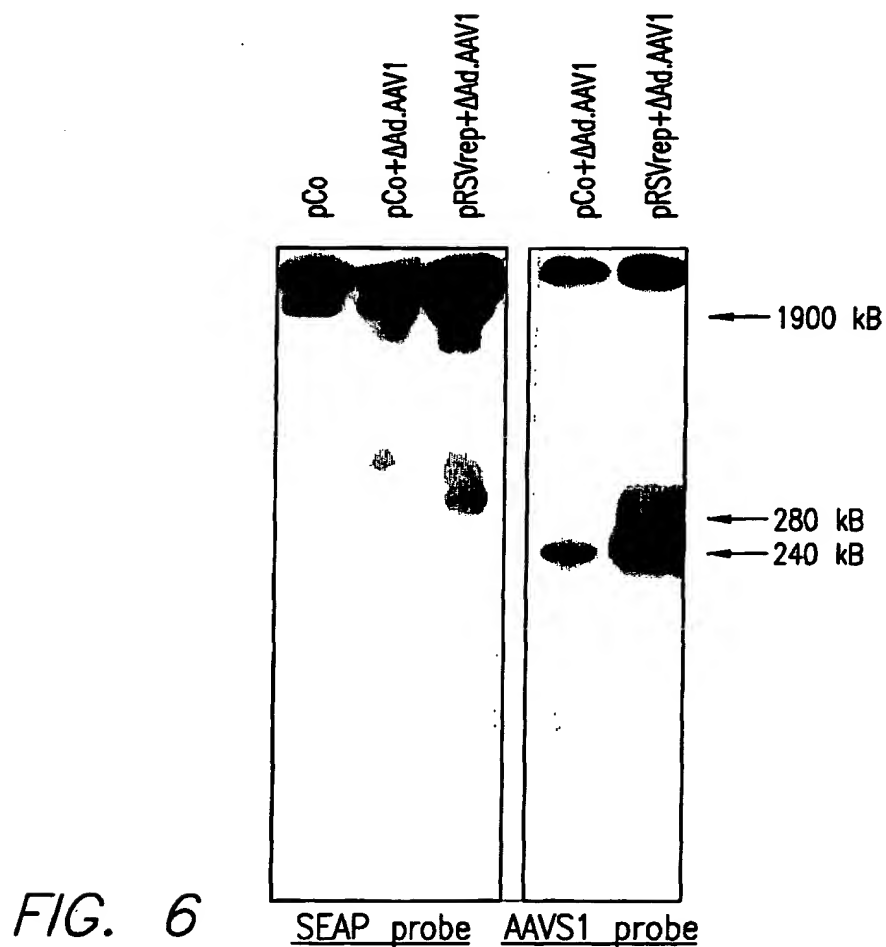


FIG. 6



6/34

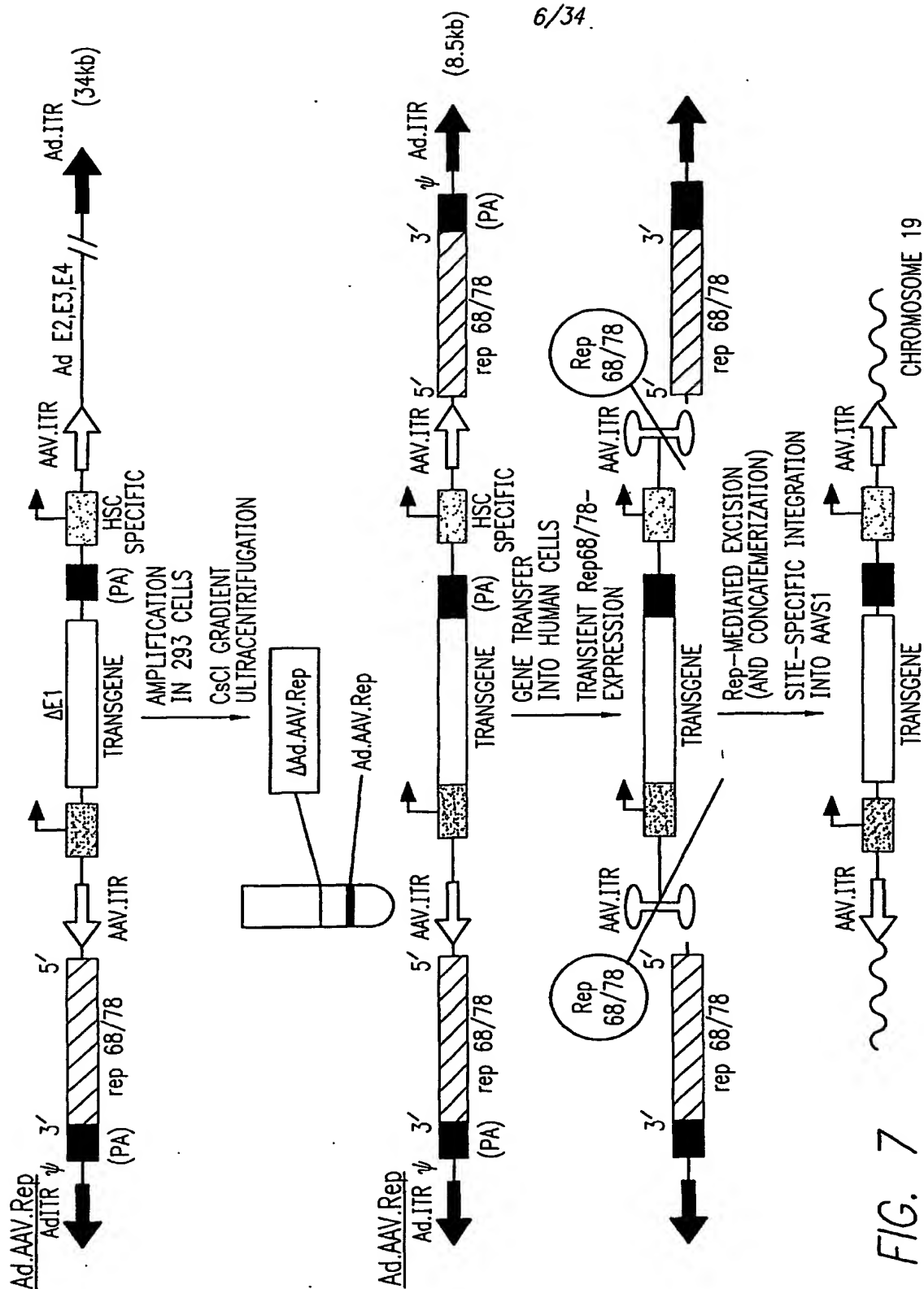


FIG. 7

7/34

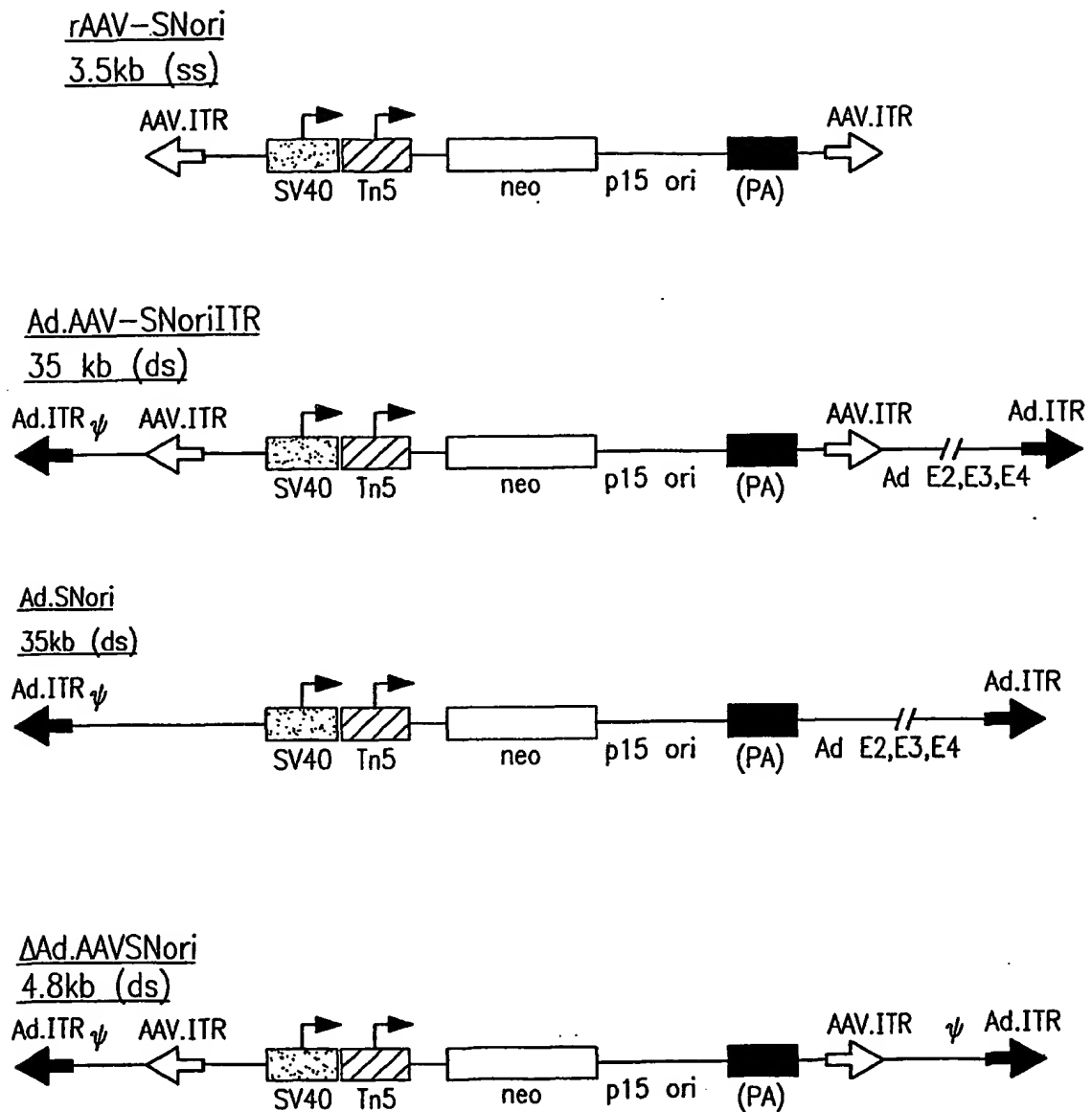
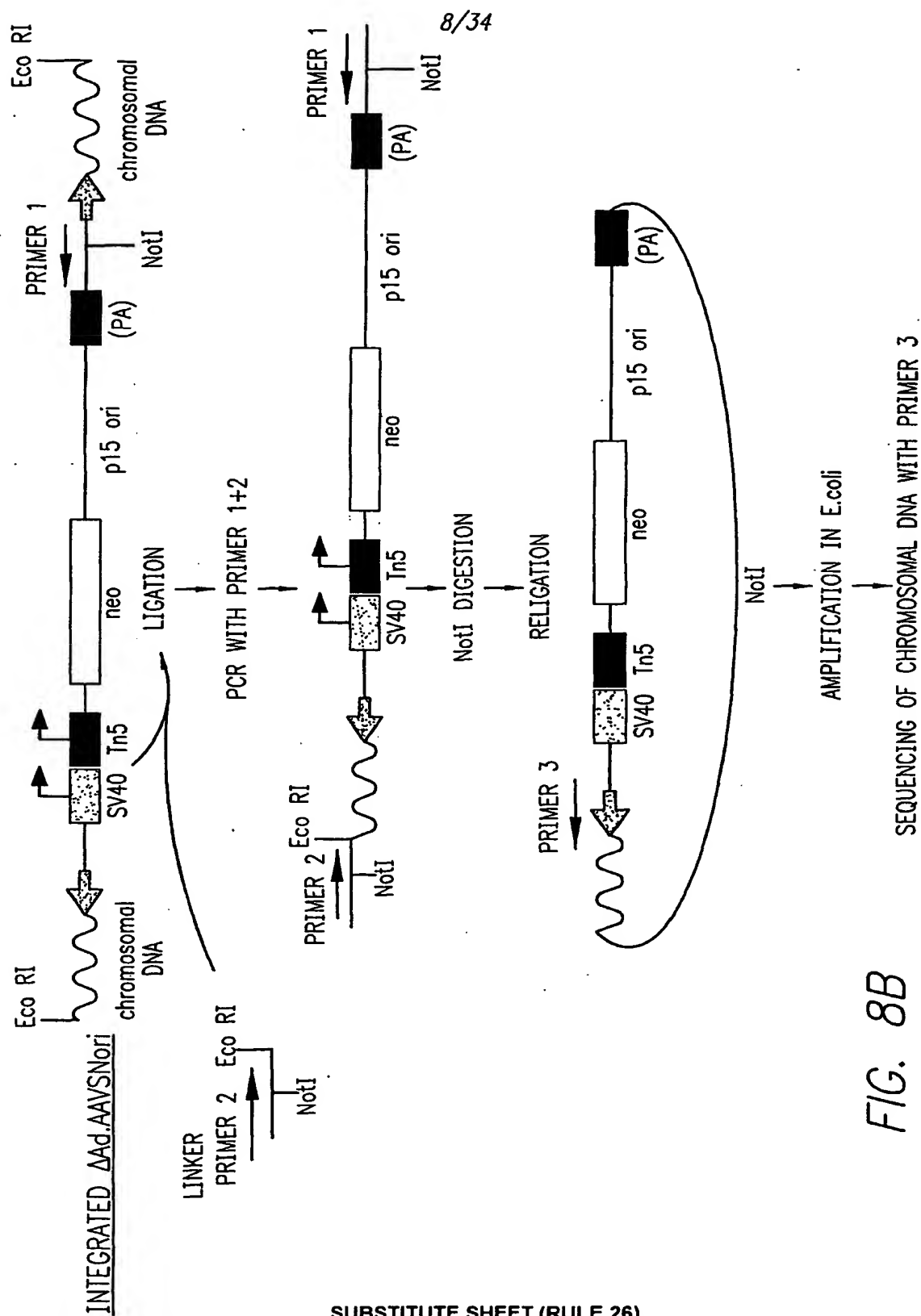


FIG. 8A



9/34

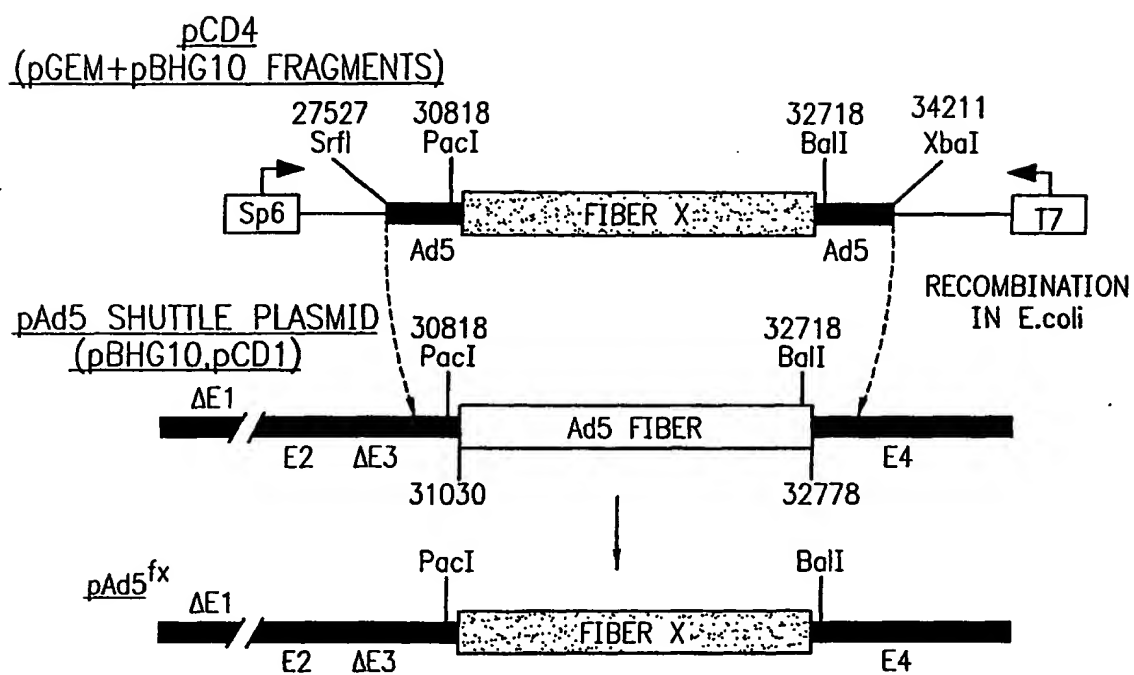


FIG. 9

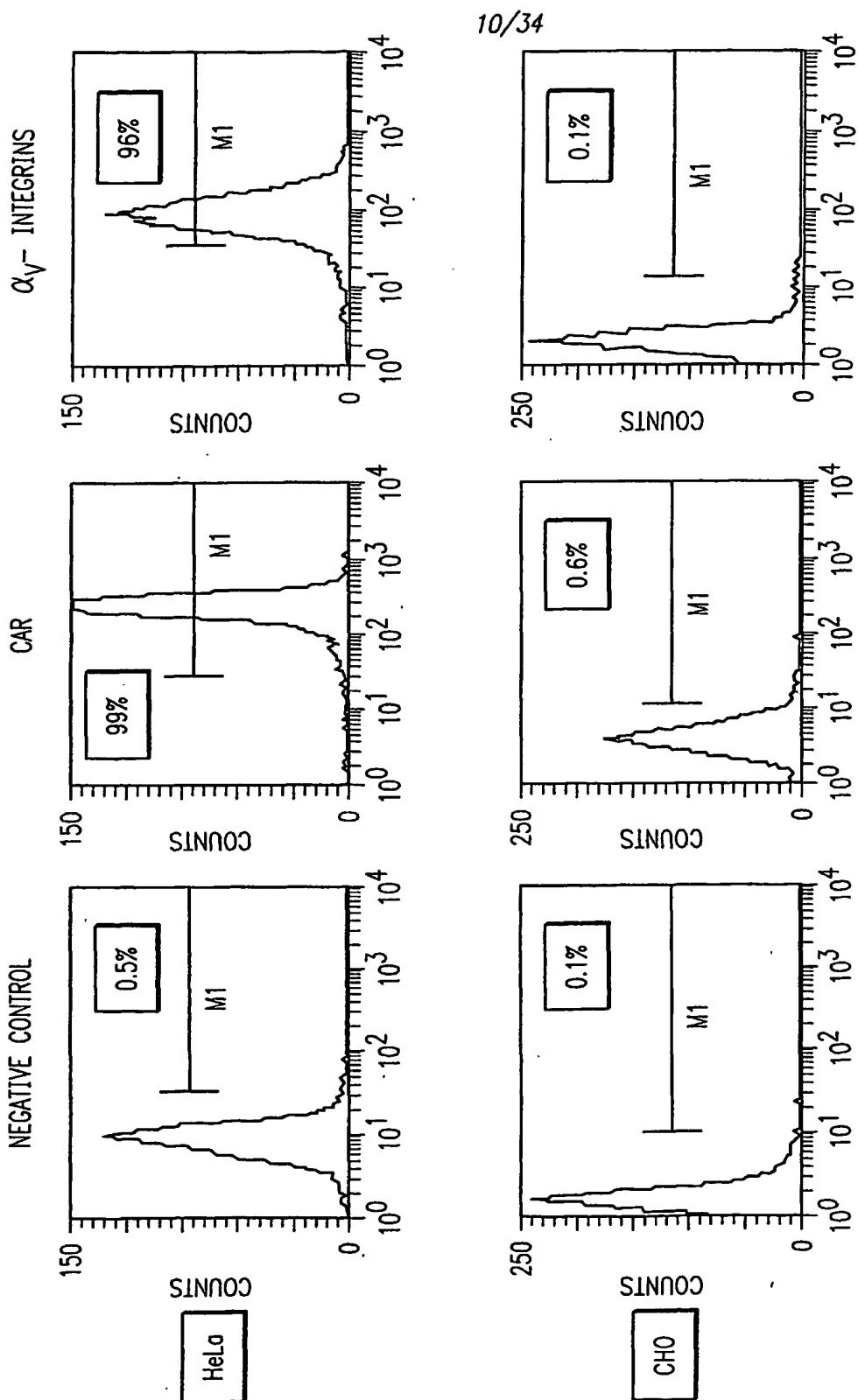


FIG. 10-i

11/34

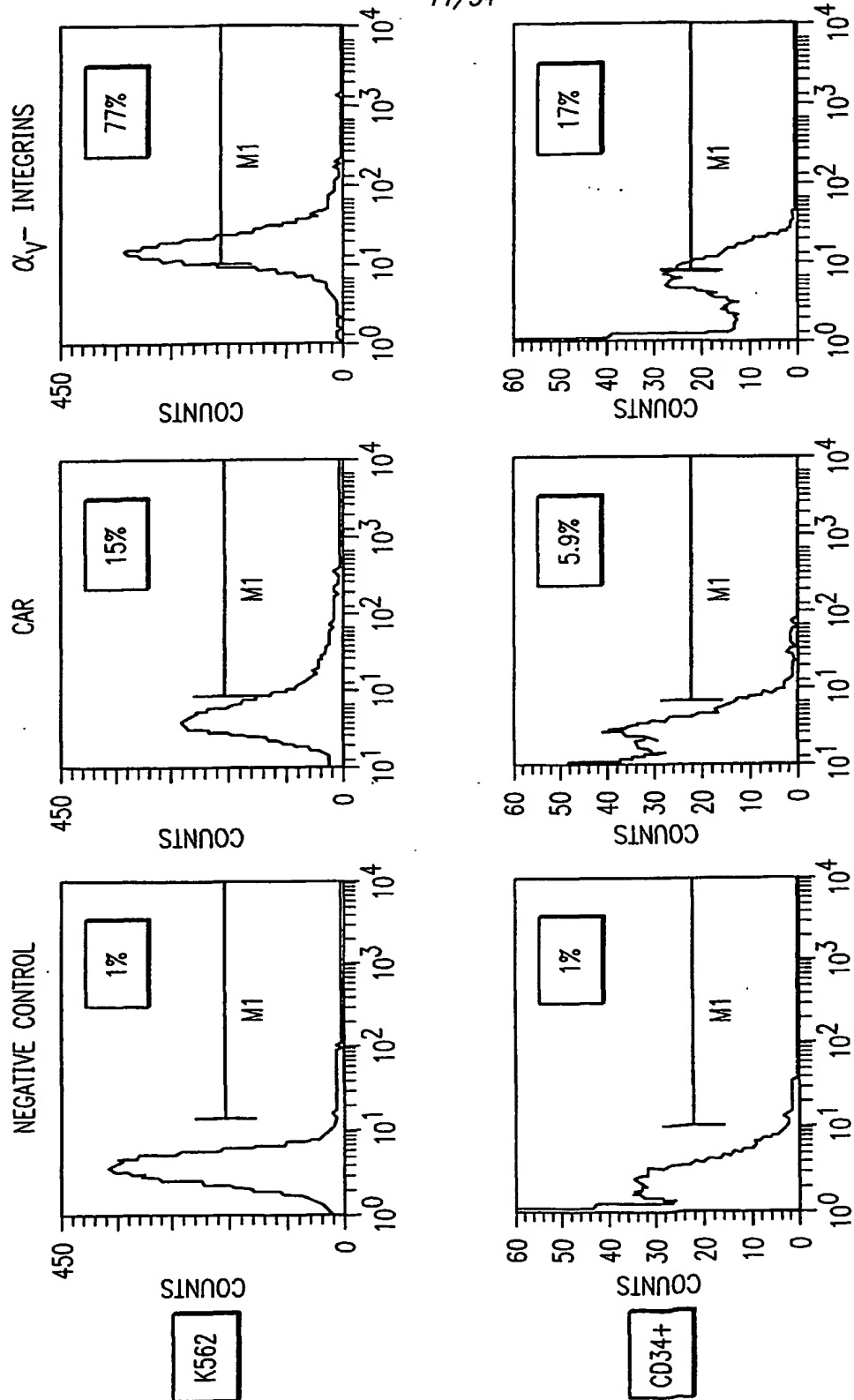
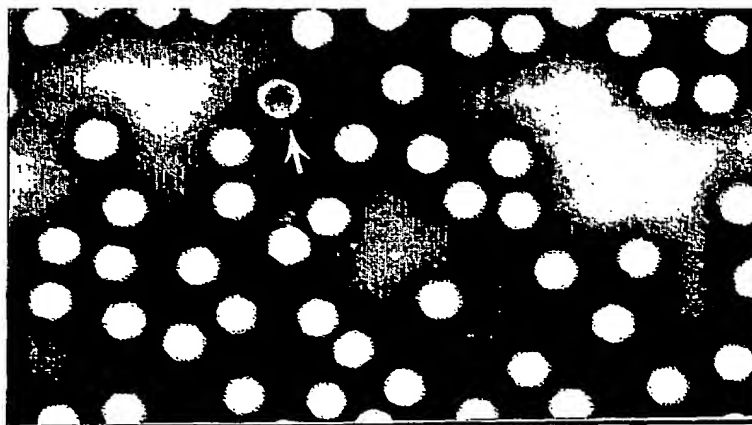


FIG. 10-ii

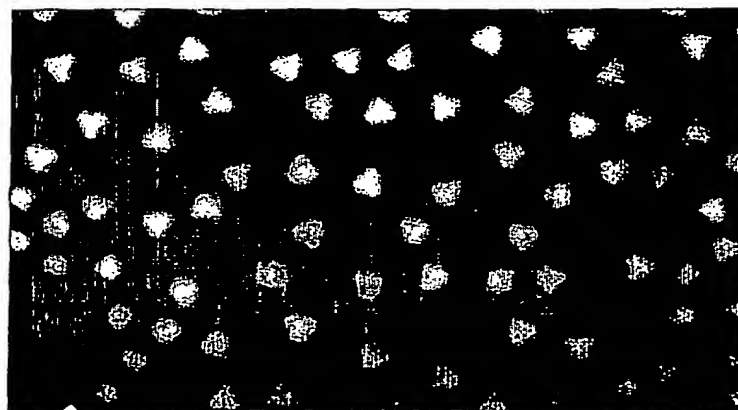
12/34



Ad 35



Ad 9



Ad 5

FIG. 11

13/34

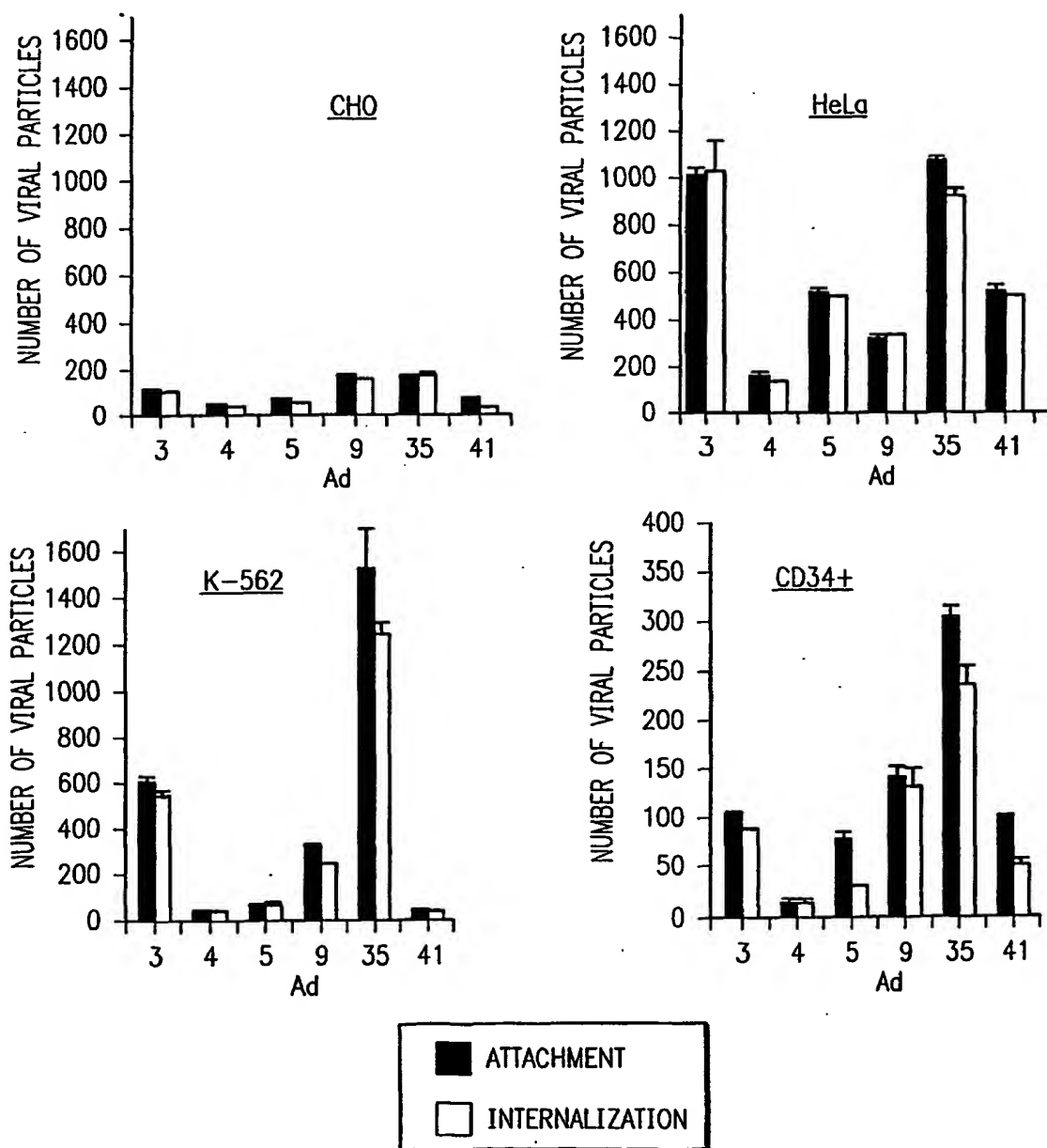
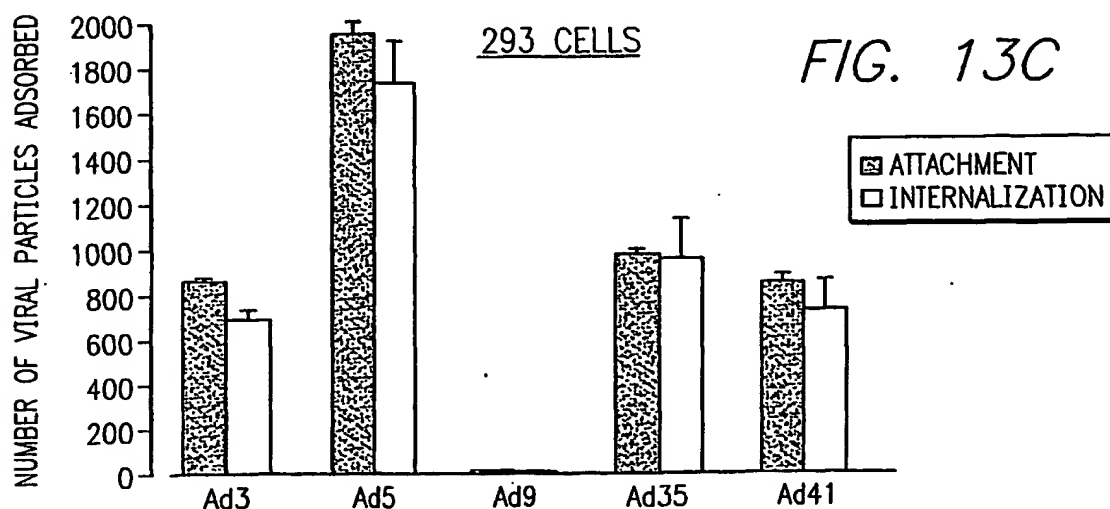
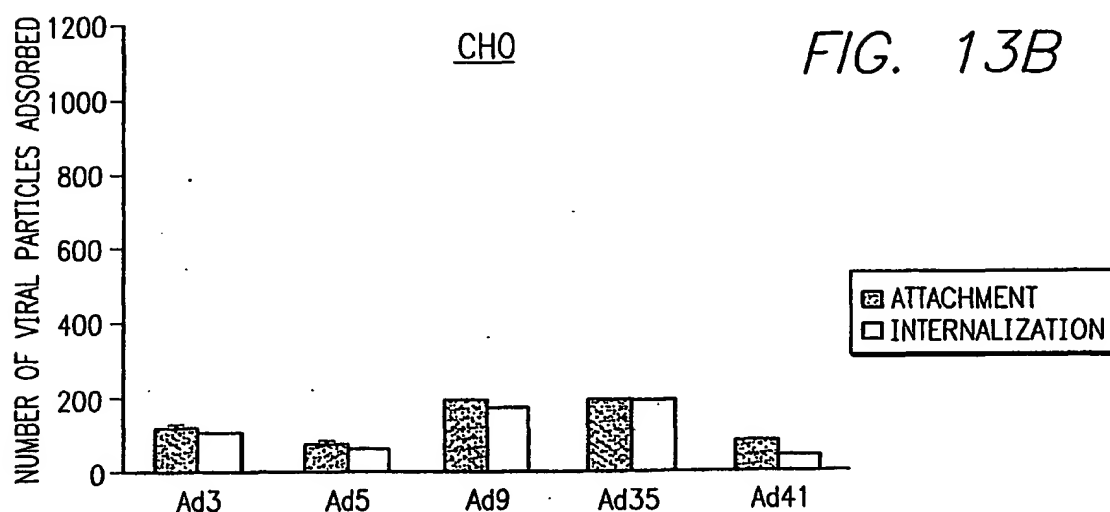
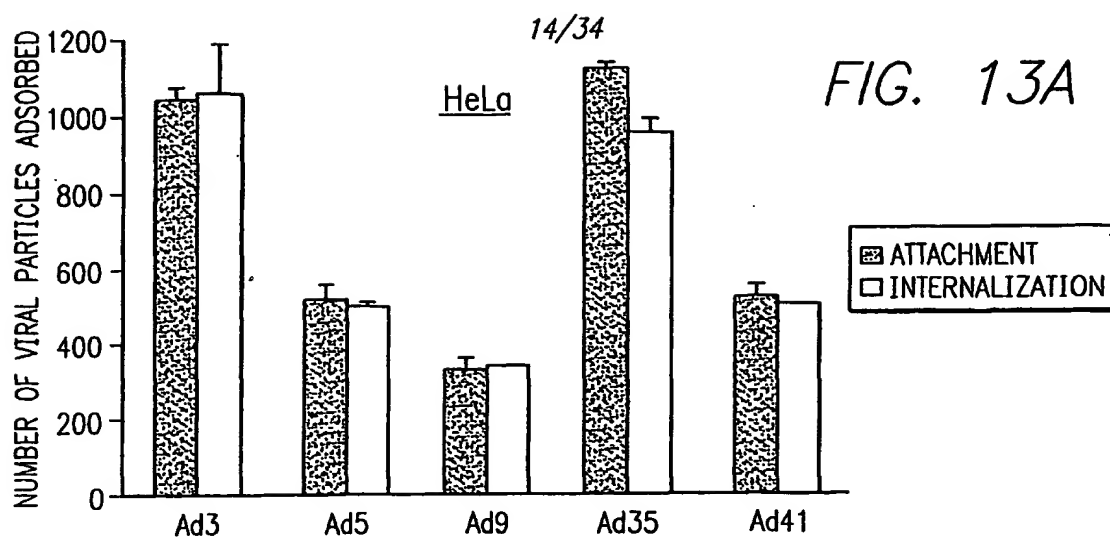


FIG. 12





15/34

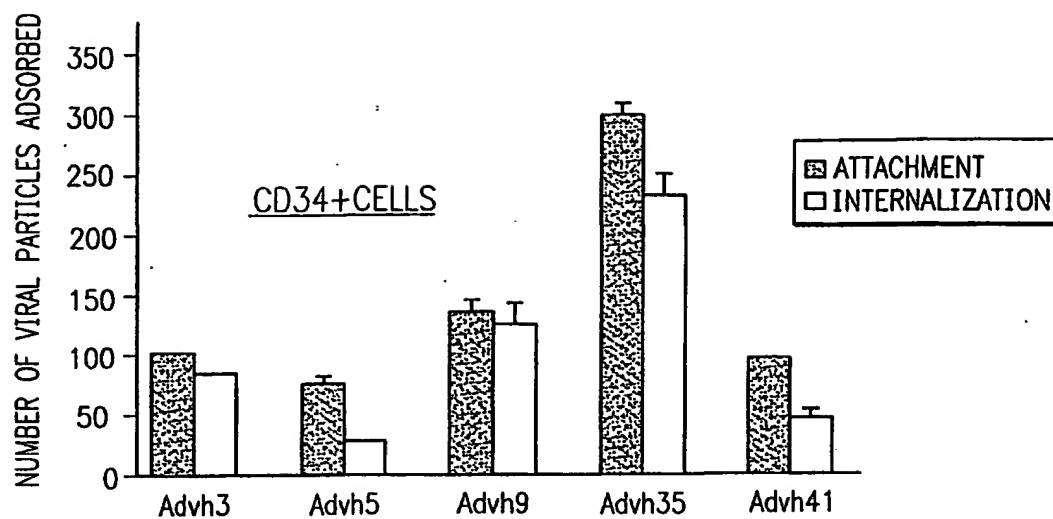


FIG. 14A

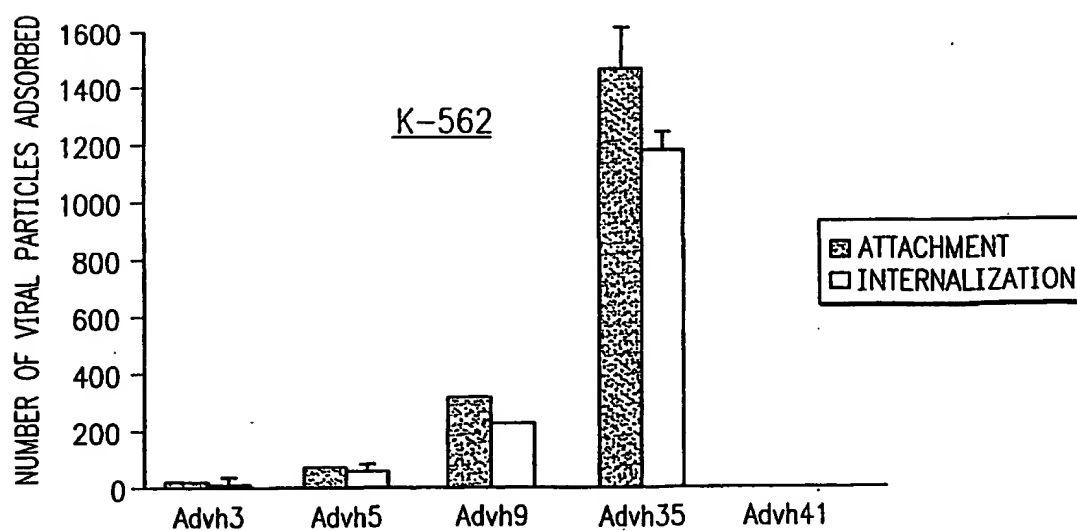


FIG. 14B

16/34

K562 cells

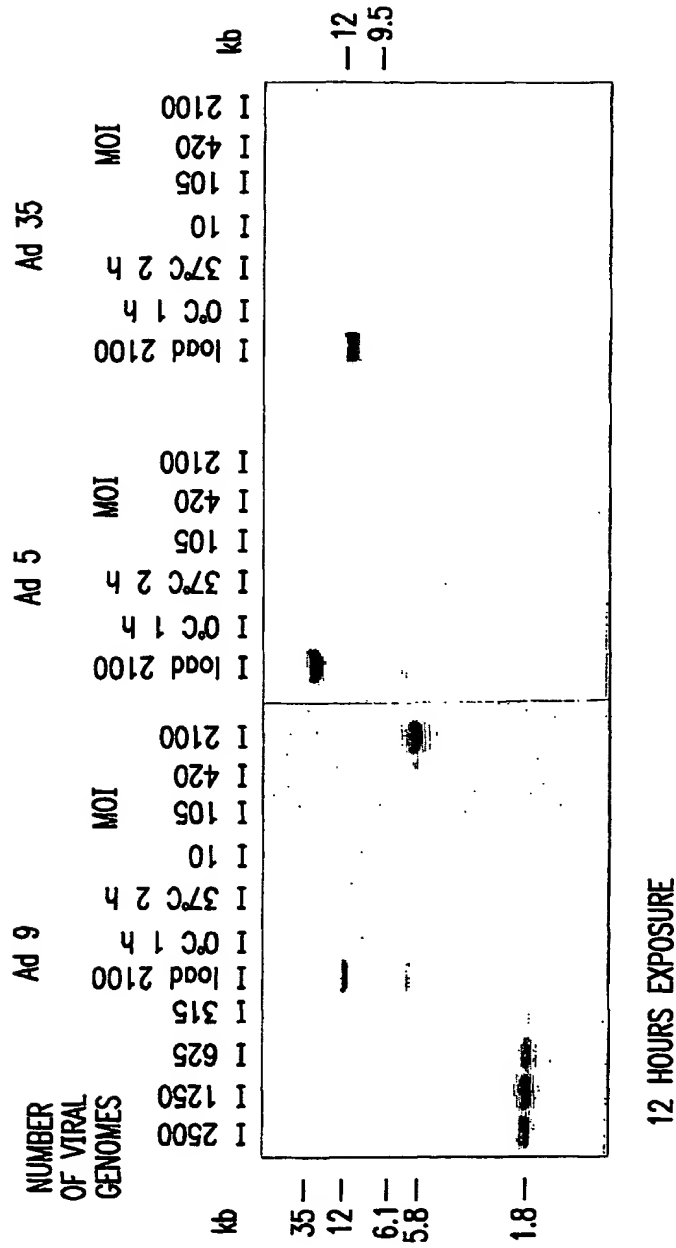


FIG. 15A

17/34

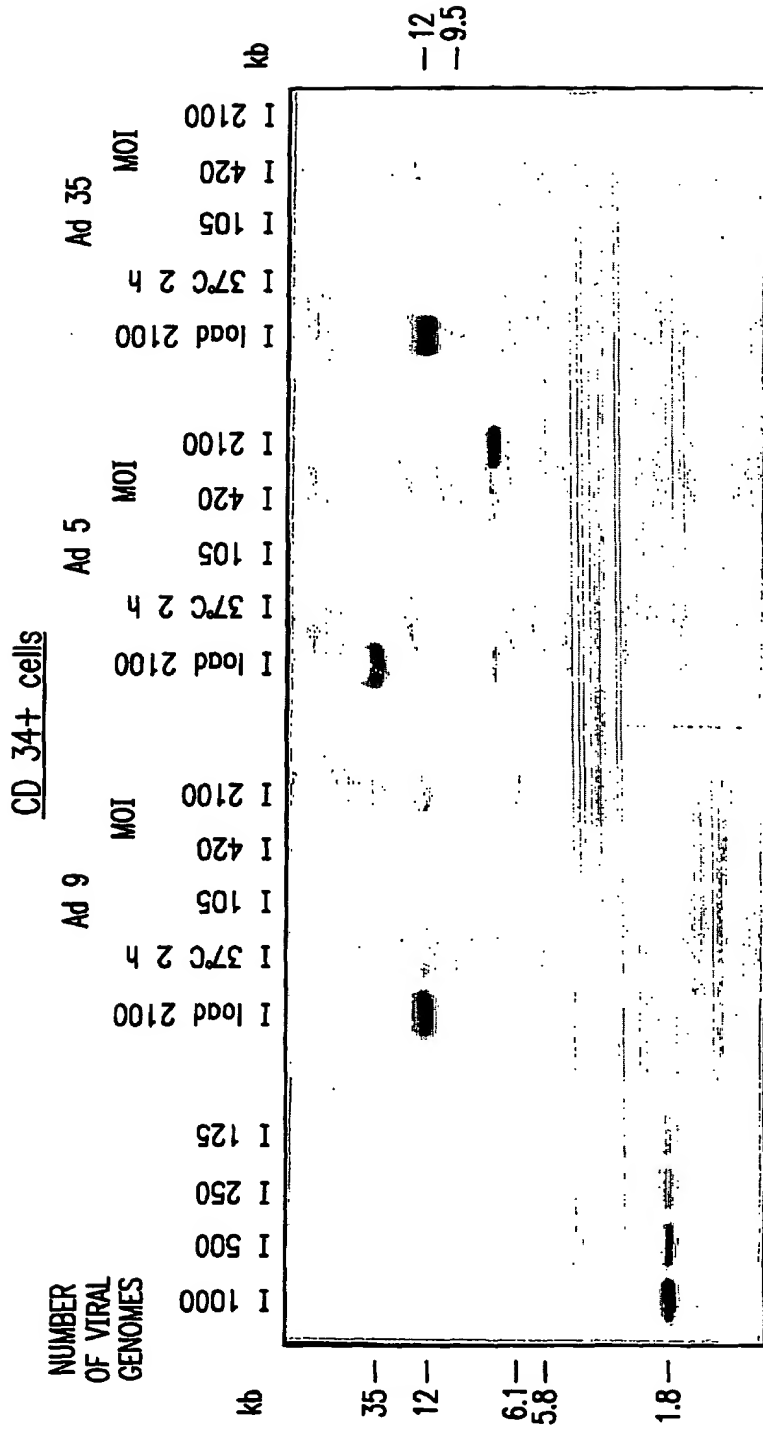


FIG. 15B

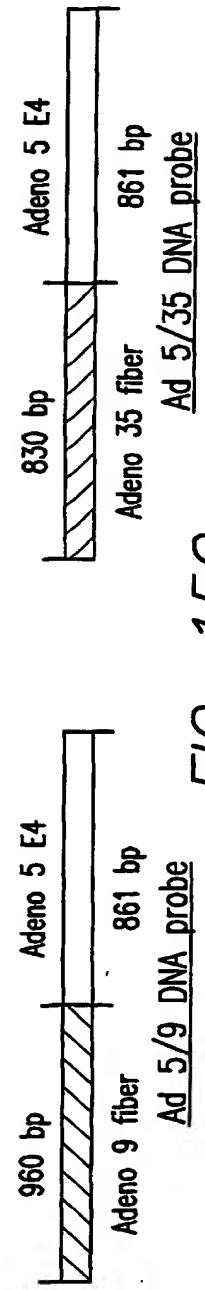


FIG. 15C

18/34

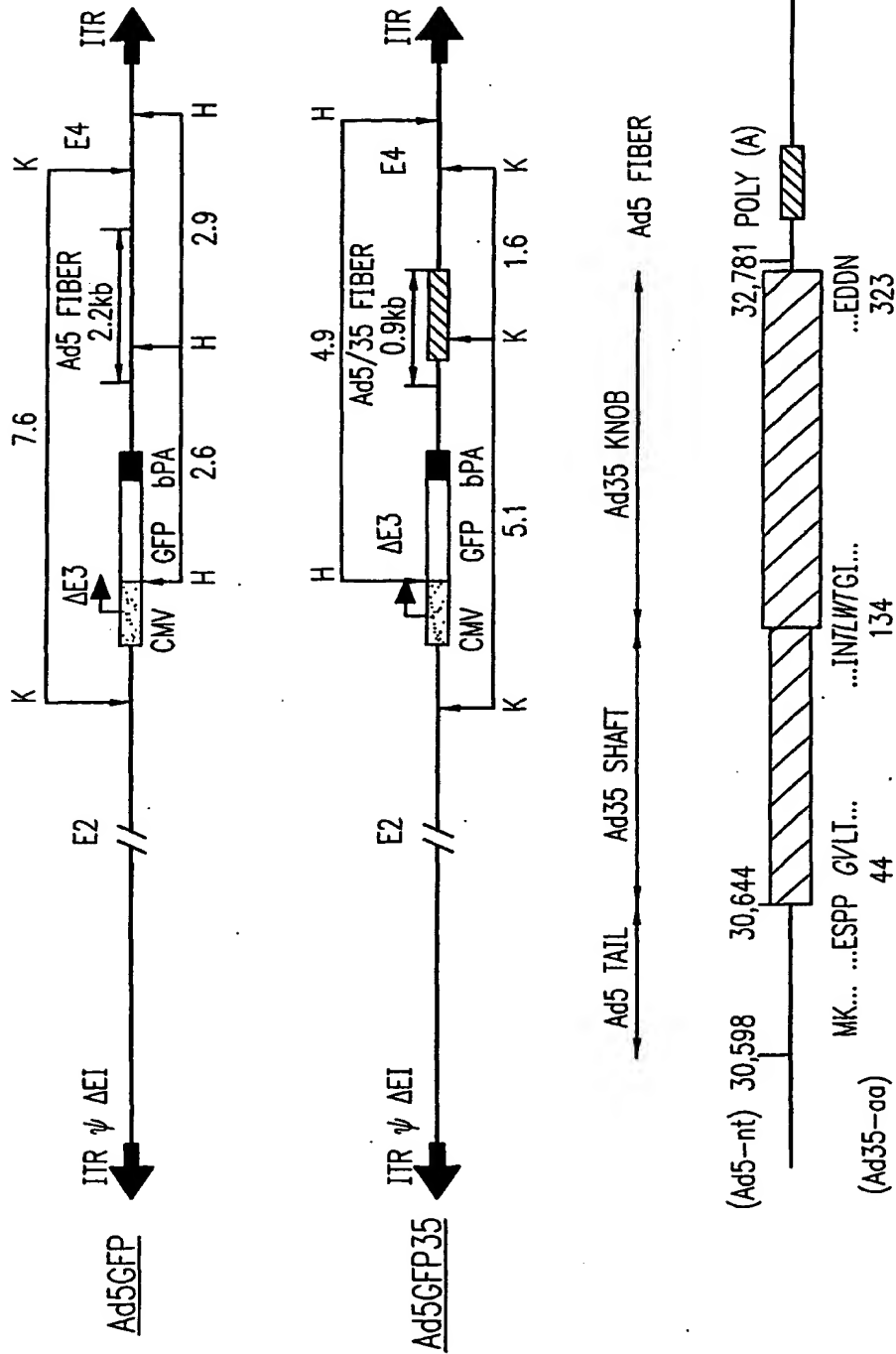


FIG. 16A

19/34

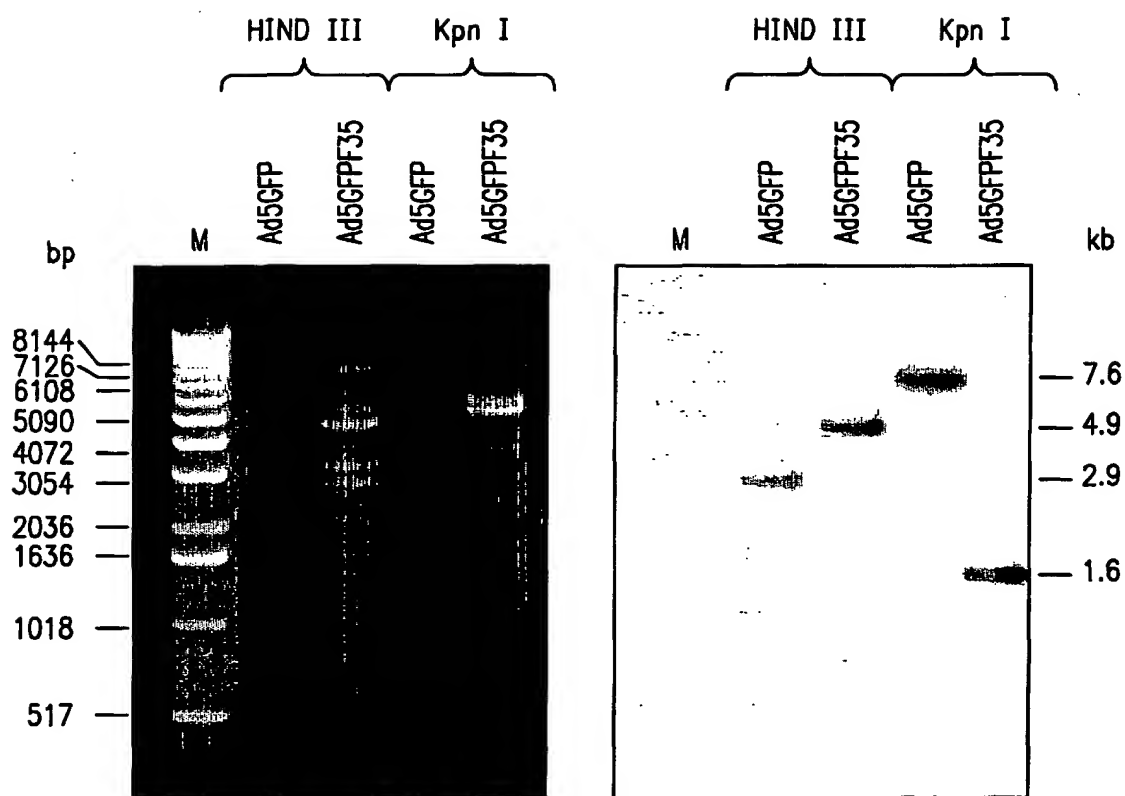


FIG. 16B

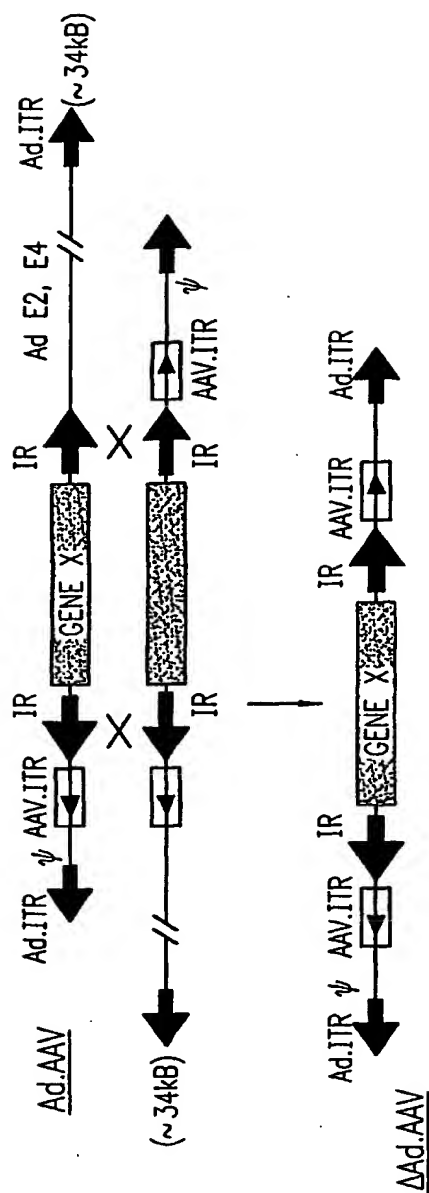


FIG. 17A

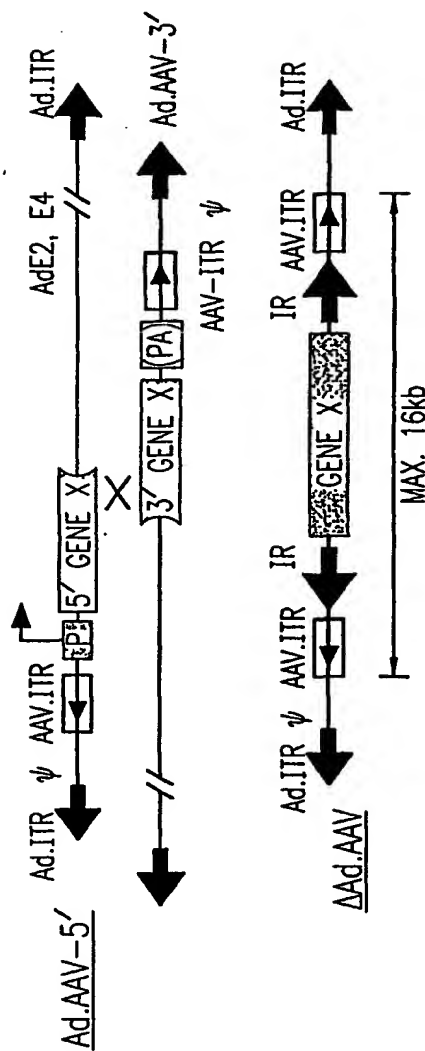


FIG. 17B

21/34

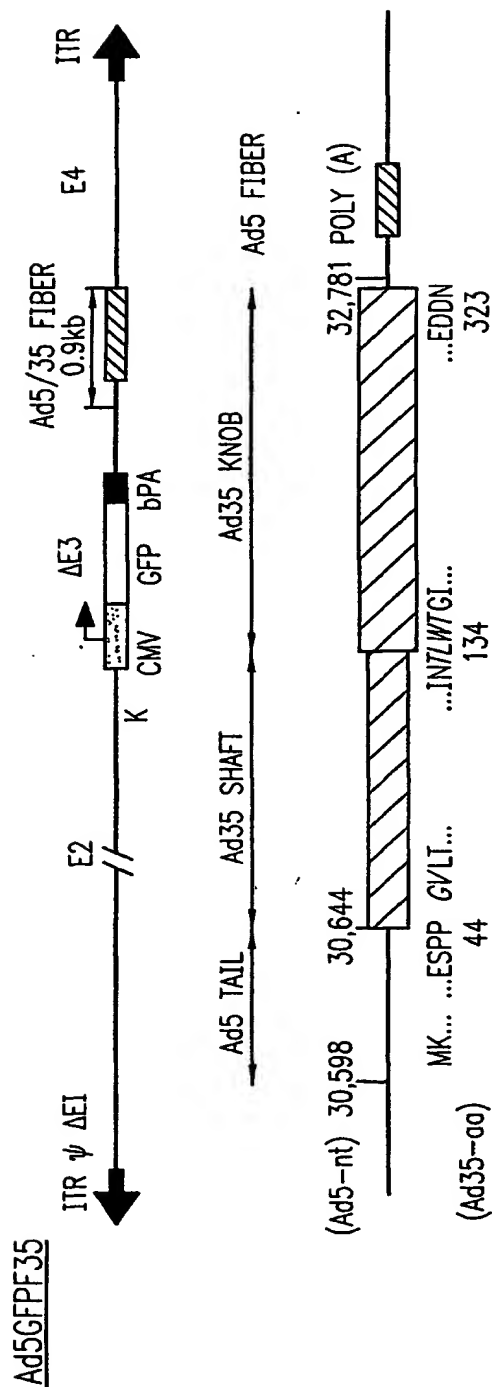


FIG. 18



22/34

■ CONTROL    □ Ad5GFP    ▨ Ad5GFPF35    ▩ Ad 35    ▪ Ad 3

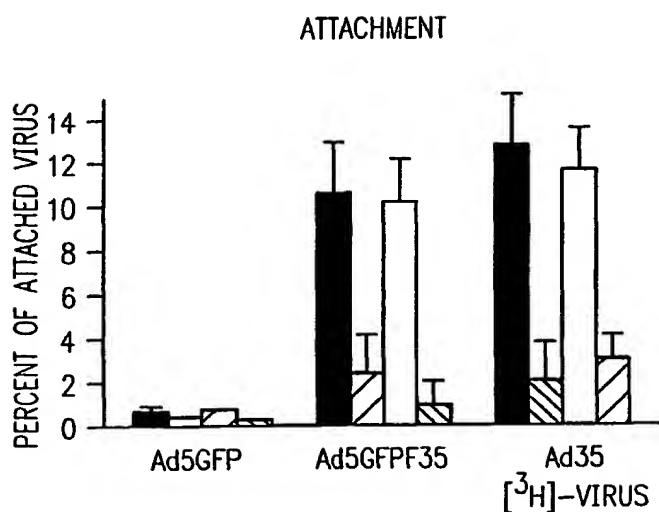


FIG. 19A

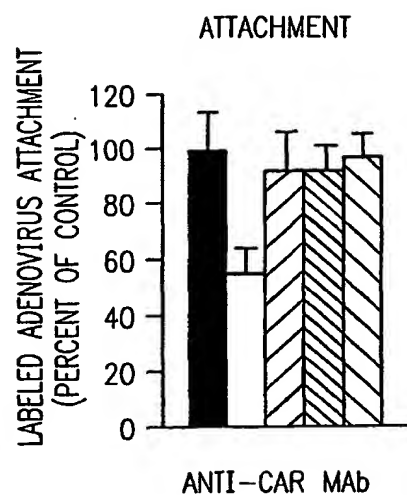


FIG. 19B

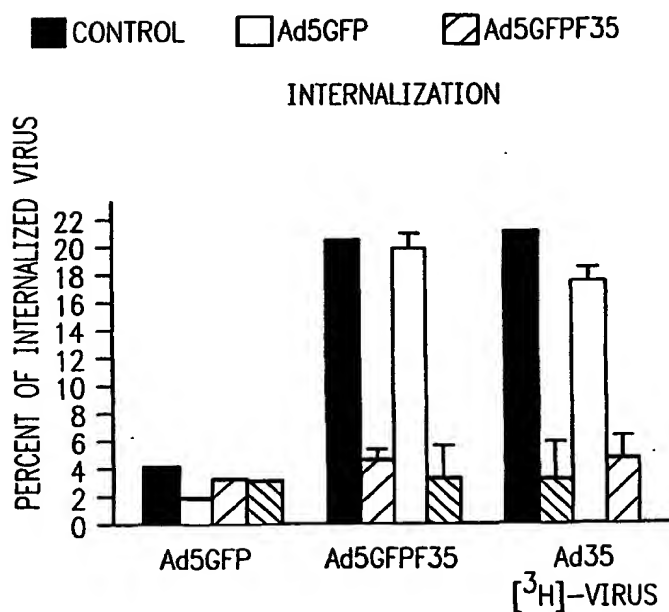


FIG. 19C

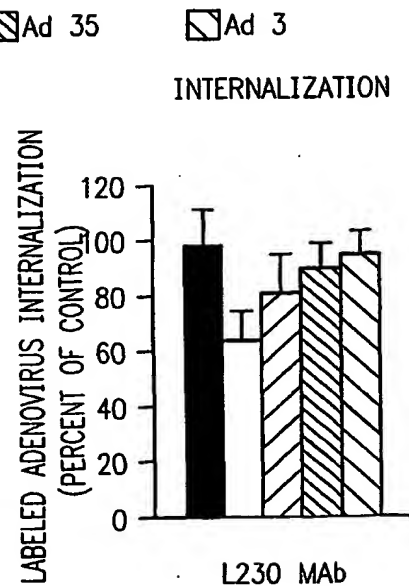


FIG. 19D

23/34

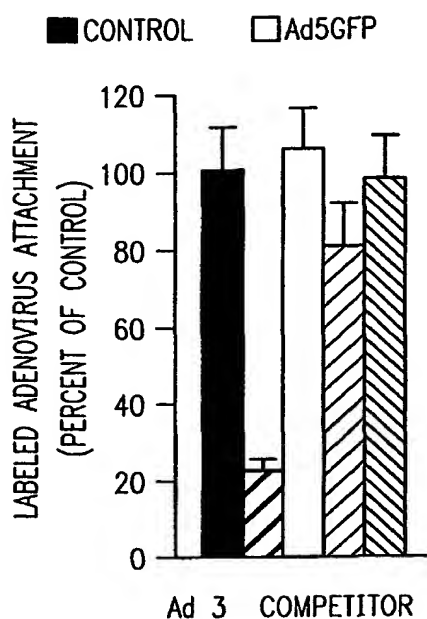


FIG. 20A

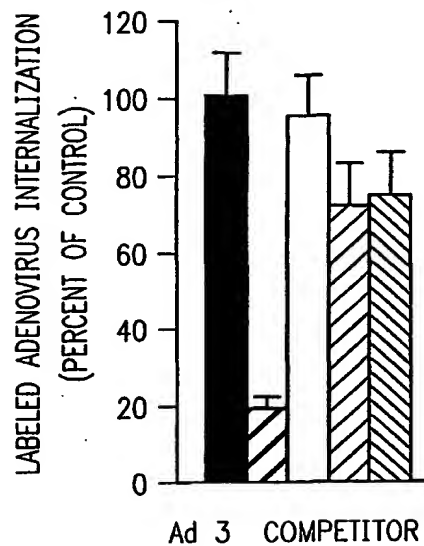


FIG. 20B

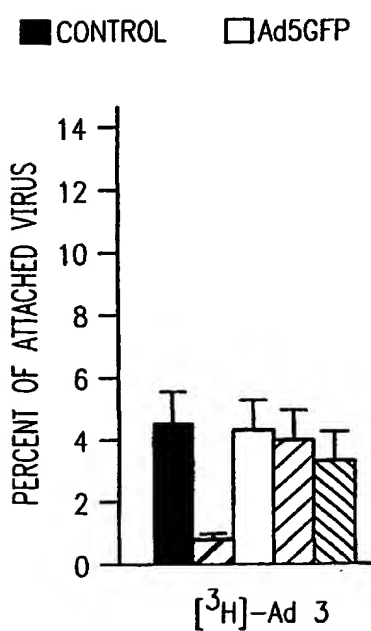


FIG. 20C

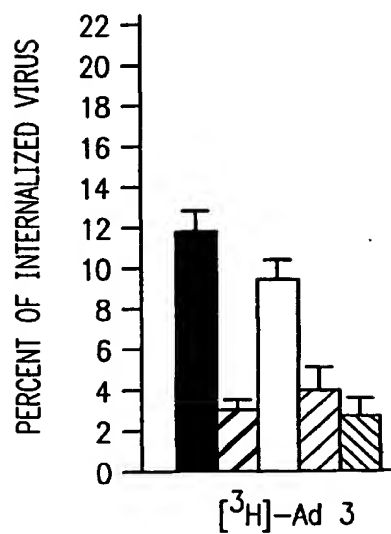


FIG. 20D

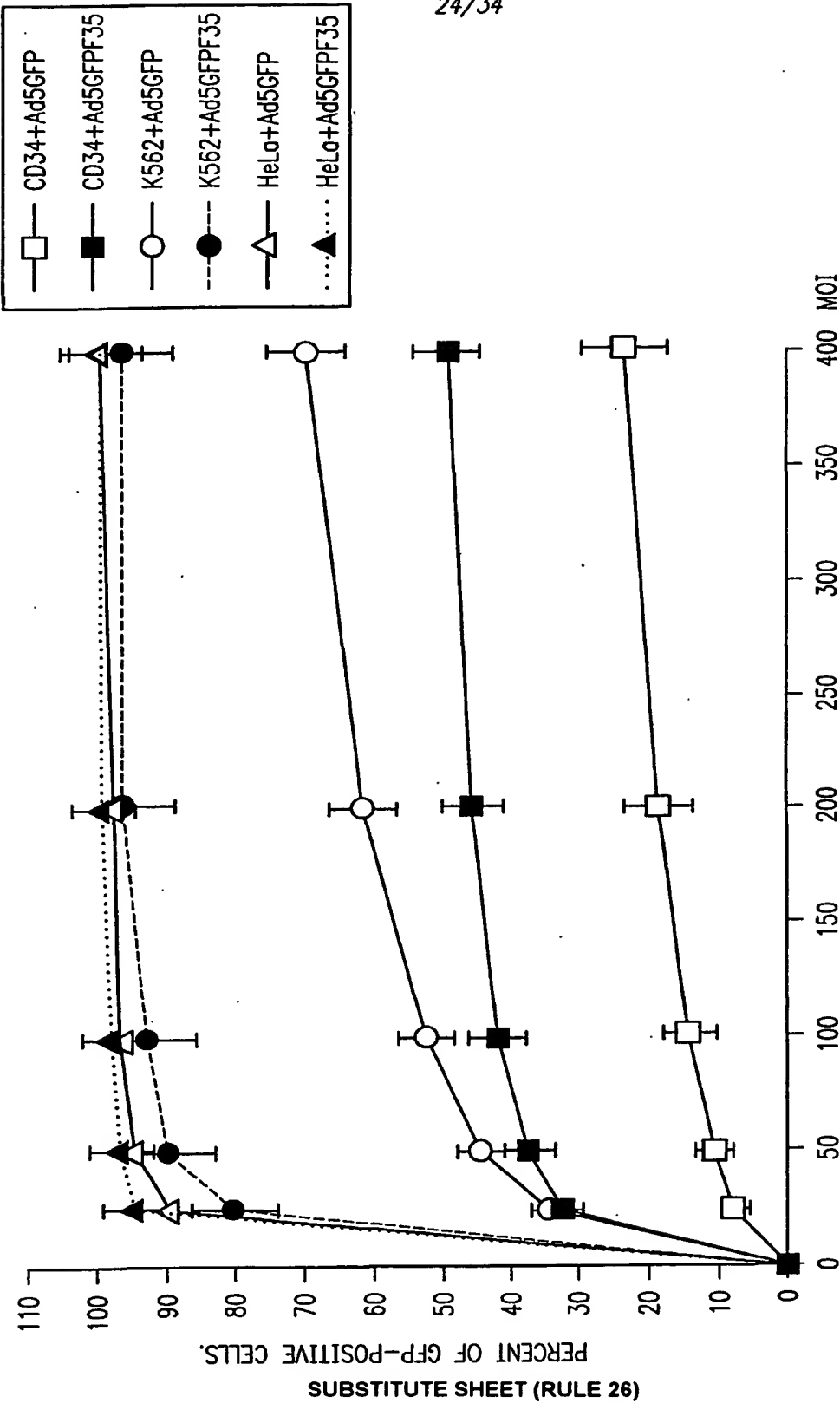


FIG. 21

25/34

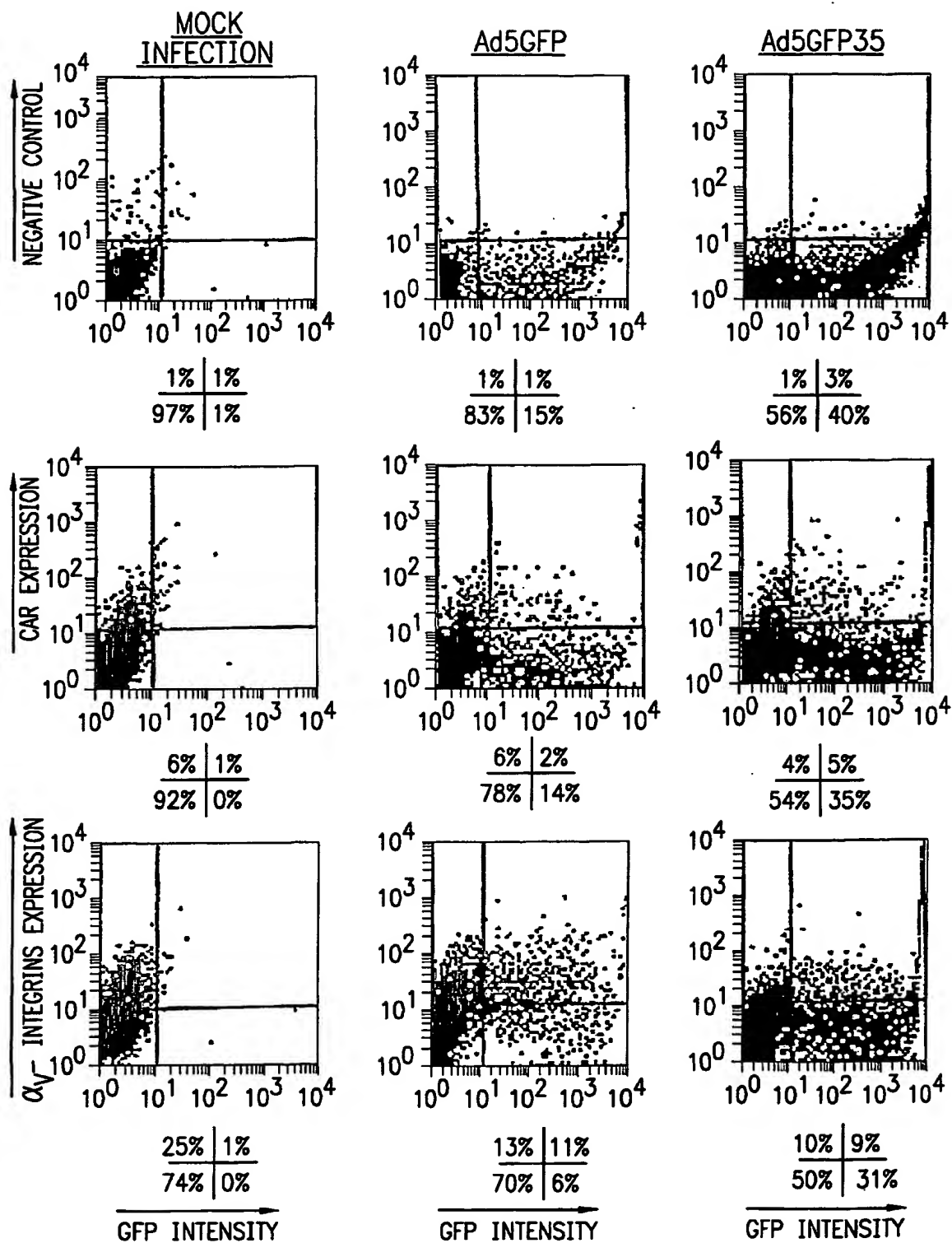


FIG. 22

SUBSTITUTE SHEET (RULE 26)

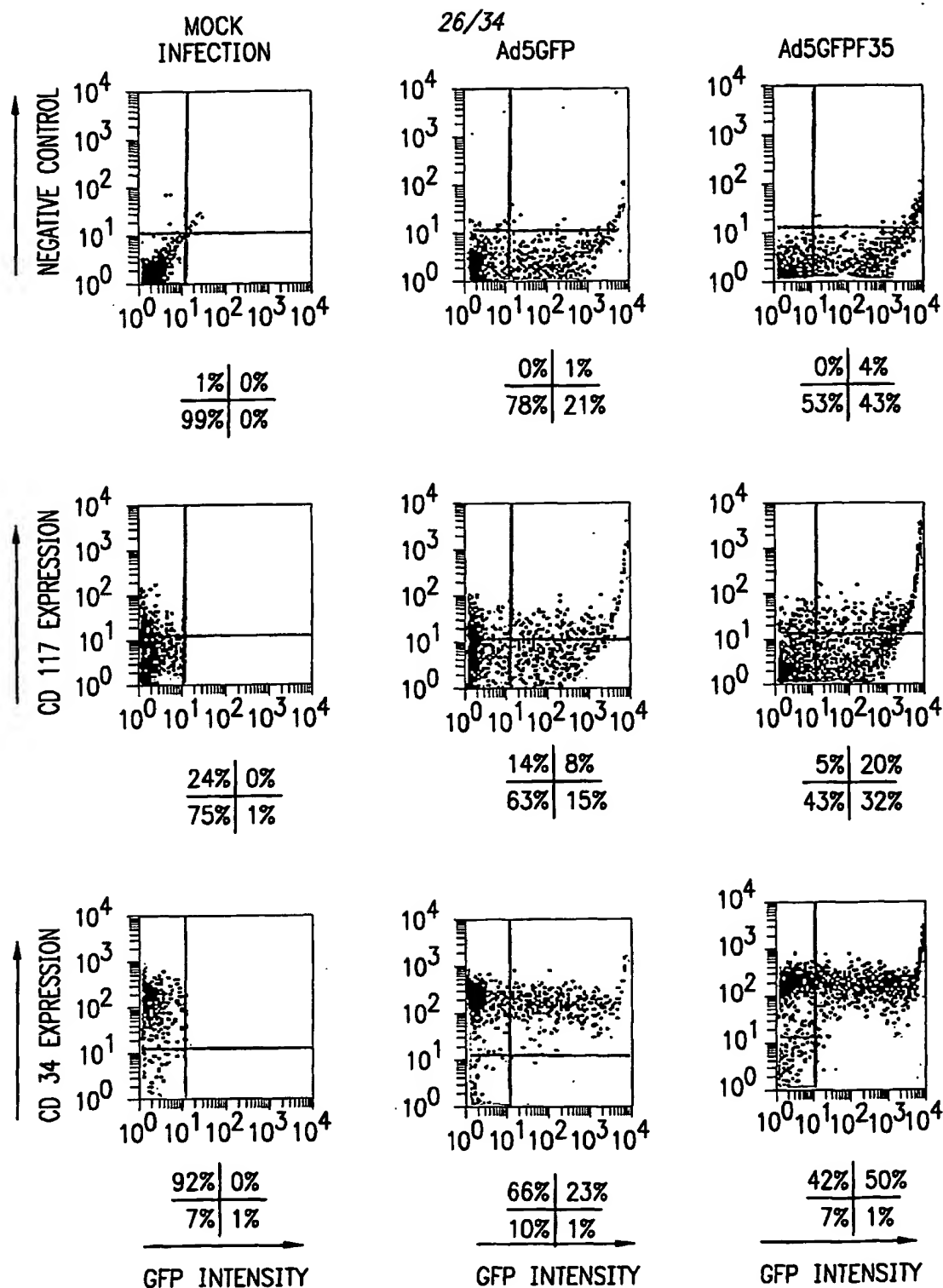


FIG. 23A

27/34

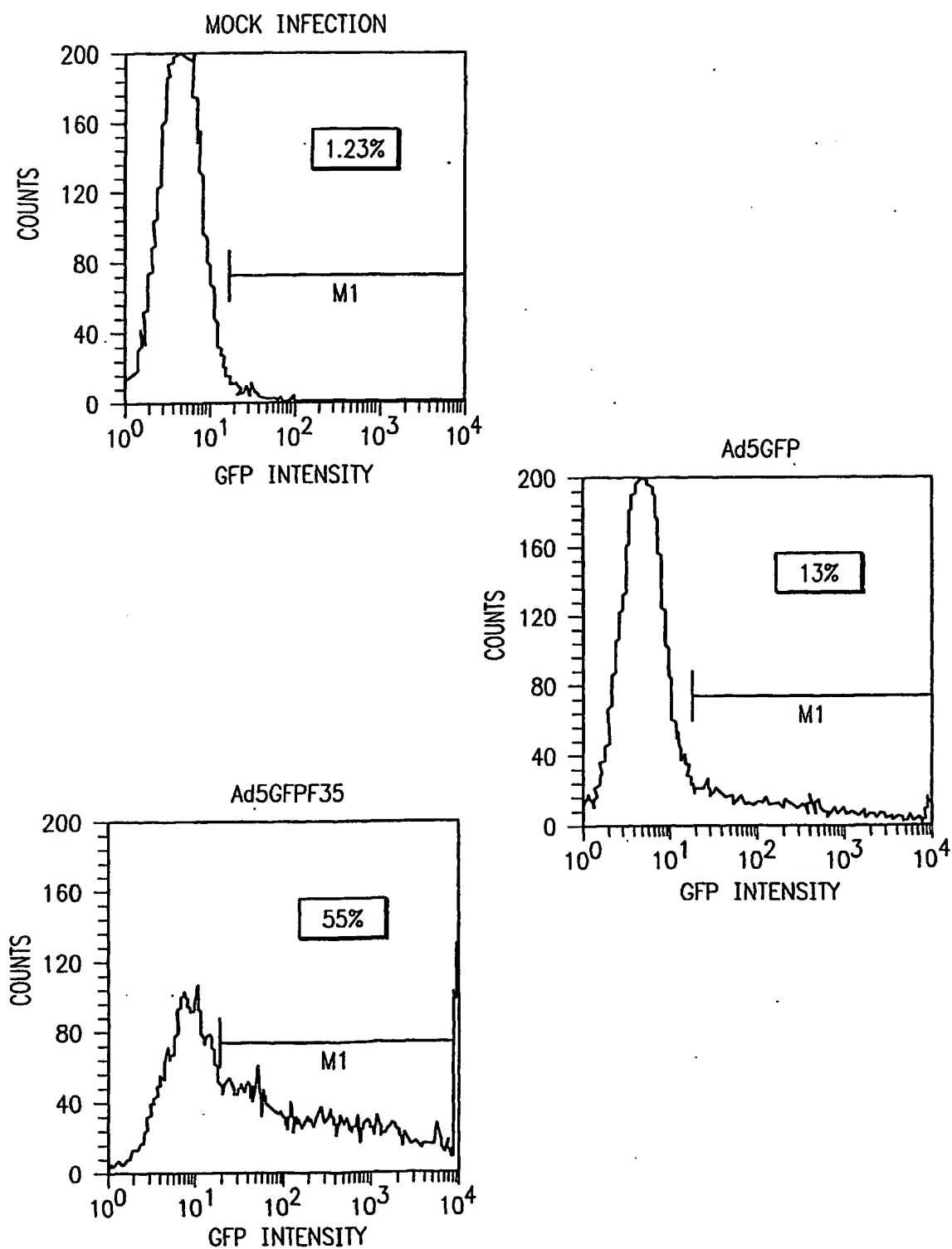
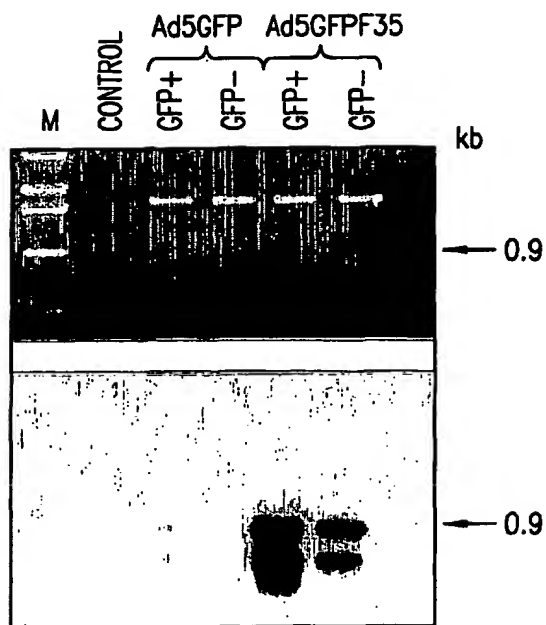
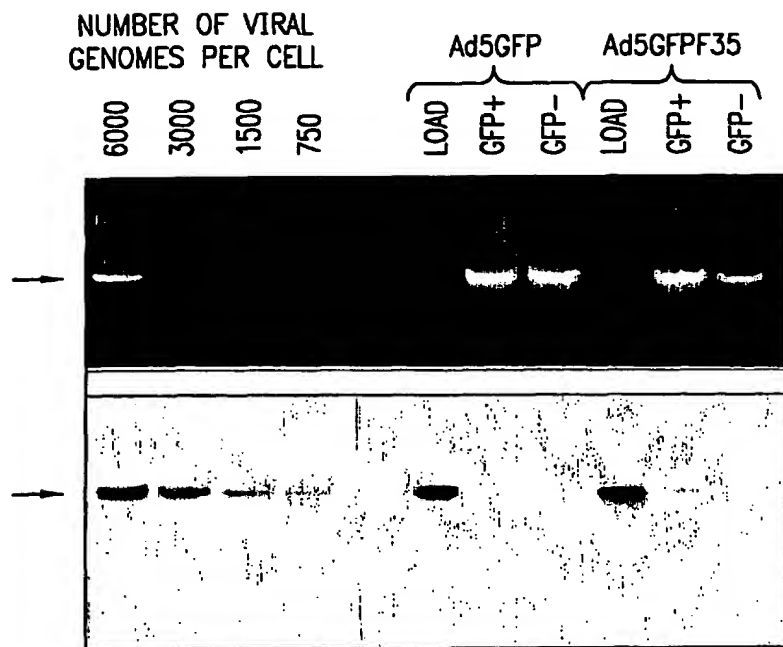


FIG. 23B

SUBSTITUTE SHEET (RULE 26)

28/34



29/34

STRUCTURE AND FEATURES OF CHIMERIC ADENOVIRUS VECTORS.



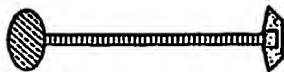
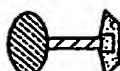
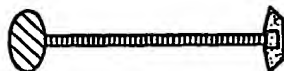

| <u>Adeno vector</u>                       |  |  |  |  |  |  |
|---|---|---|---|--|---|---|
| Binding to CAR                            | YES   | YES   | YES   | YES  | NO  | NO  |
| genome/PFU ratio                          | 13  | 143   | 29  | 582  | 17  | 14  |
| <u>Attachment to/internalization into</u> |   |   |   |  |   |   |
| 293                                       | +++   | ++  | ++  | +/---  | +++   | +++   |
| K562                                      | -/+   | --/+  | --/+  | ---  | +++   | +++   |
| Y79                                       | ++  | +   | +   | +/+  | +   | +   |
| <u>Gene transfer efficiency</u>           |   |   |   |  |   |   |
| 293                                       | +++   | ++  | ++  | +/---  | +++   | +++   |
| K562                                      | -/+   | --/+  | --/+  | ---  | +++   | +++   |
| Y79                                       | ++  | +   | +   | +/+  | +   | +   |

FIG. 25



30/34

Top View of Adenovirus Type 5 Fiber Knob Domain

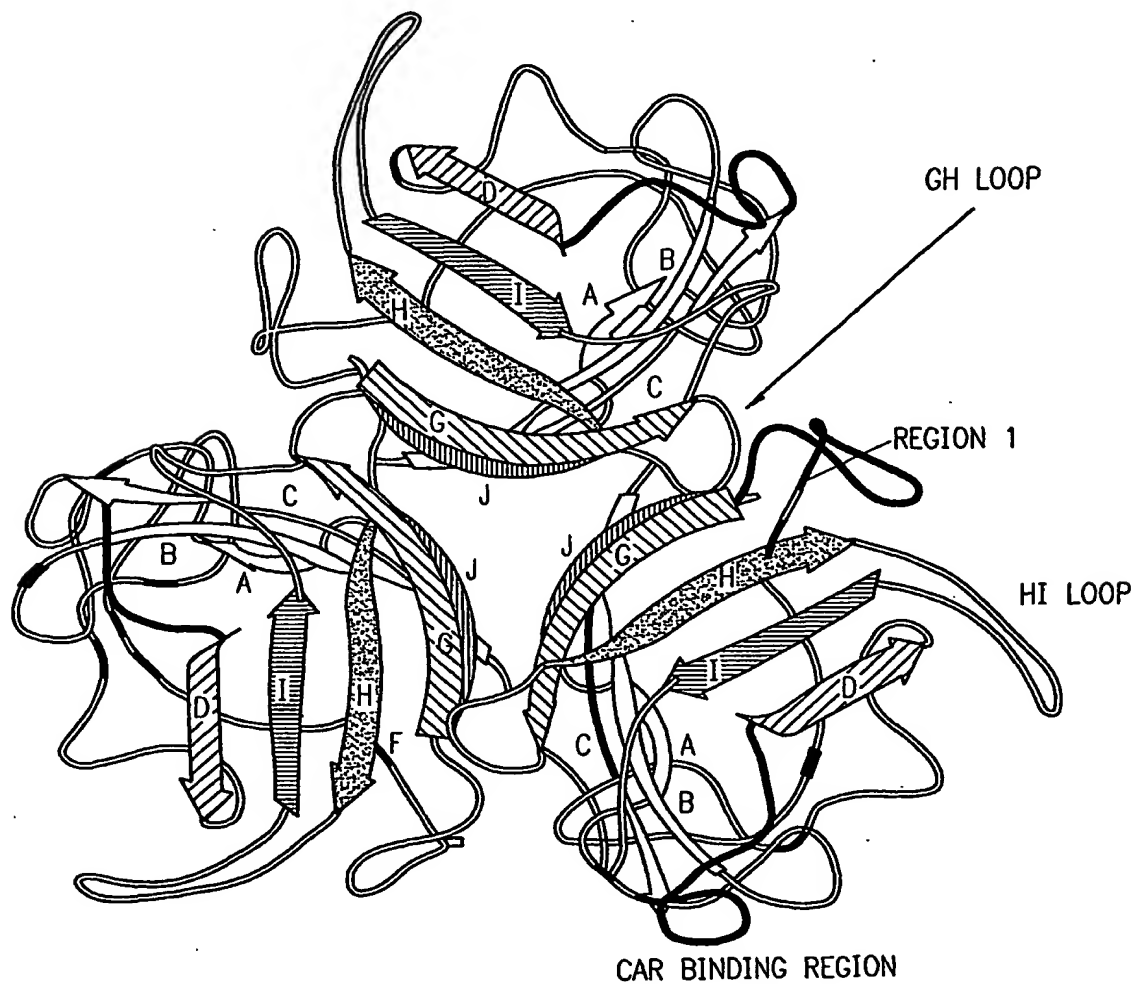


FIG. 26

Amino Acid Composition of Wt and Modified Adeno 5 GH-loop

Adeno 5 wt GH-loop      V Y] L N G D K T K [P V

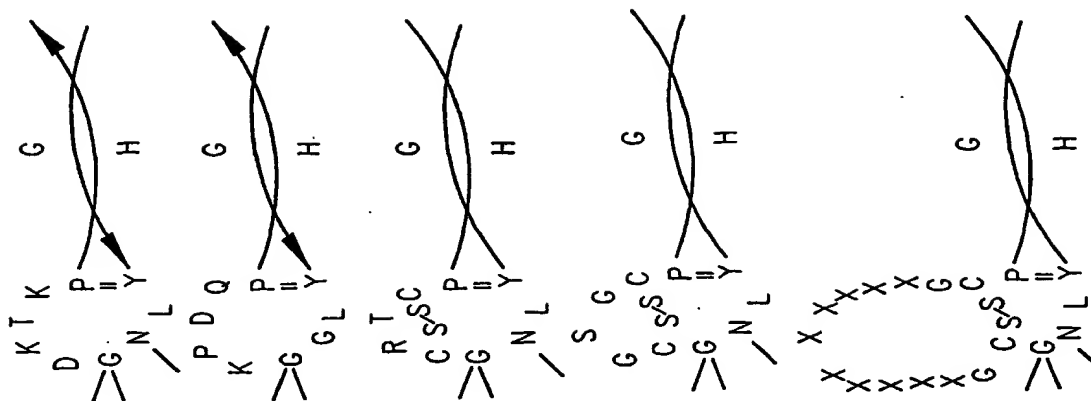
Adeno 5/9 GH-loop chimera      V Y] L G G K P D Q [P V

Adeno 5 GH-Cys 1      V Y] L N G C G S C [P V

Adeno 5 GH-Cys 2      V Y] L N G C G S G C [P V

Adeno 5 GH-peptide      V Y] L N G C G (X)<sup>10</sup> G C [P V

FIG. 27



32/34

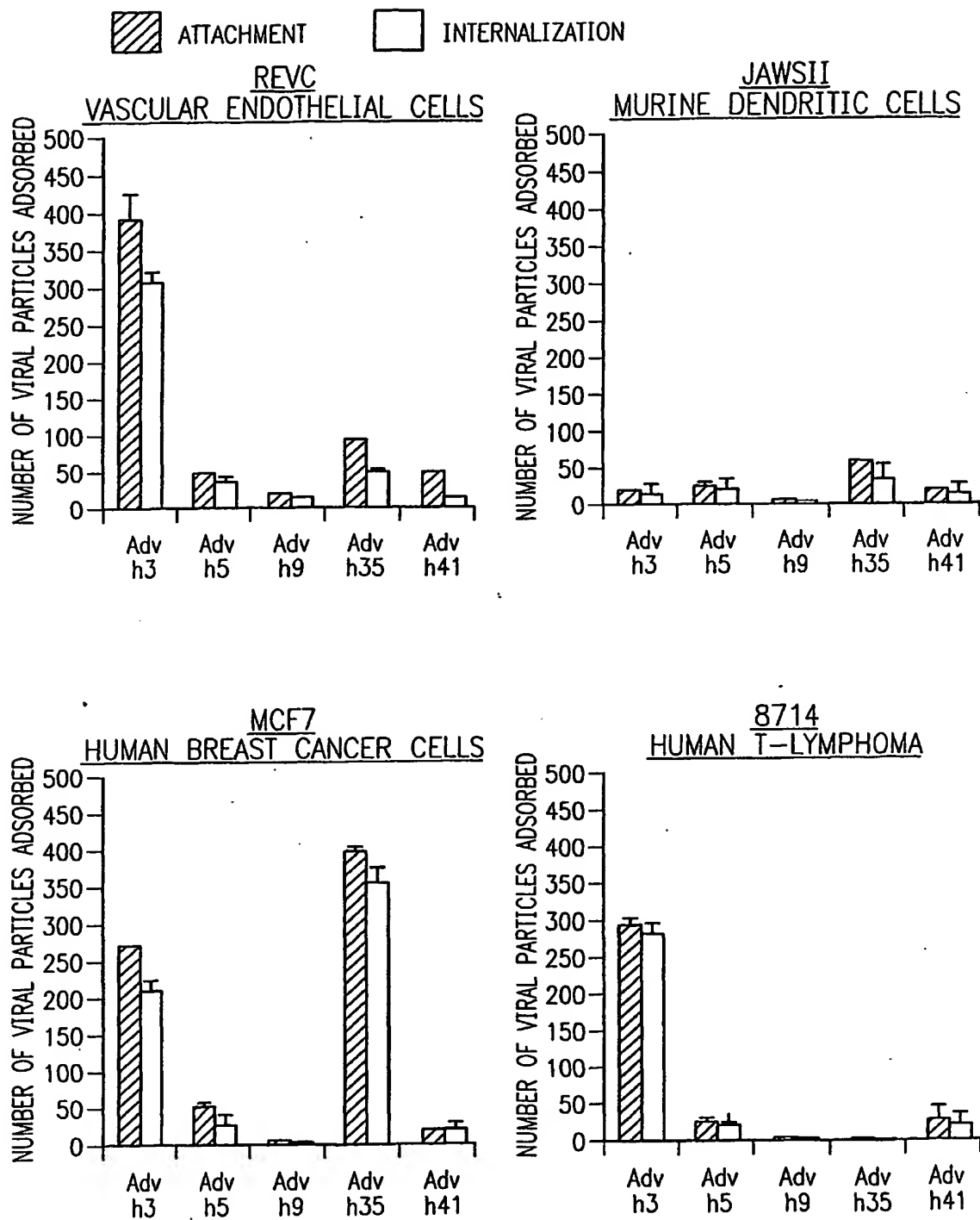


FIG. 28

33/34

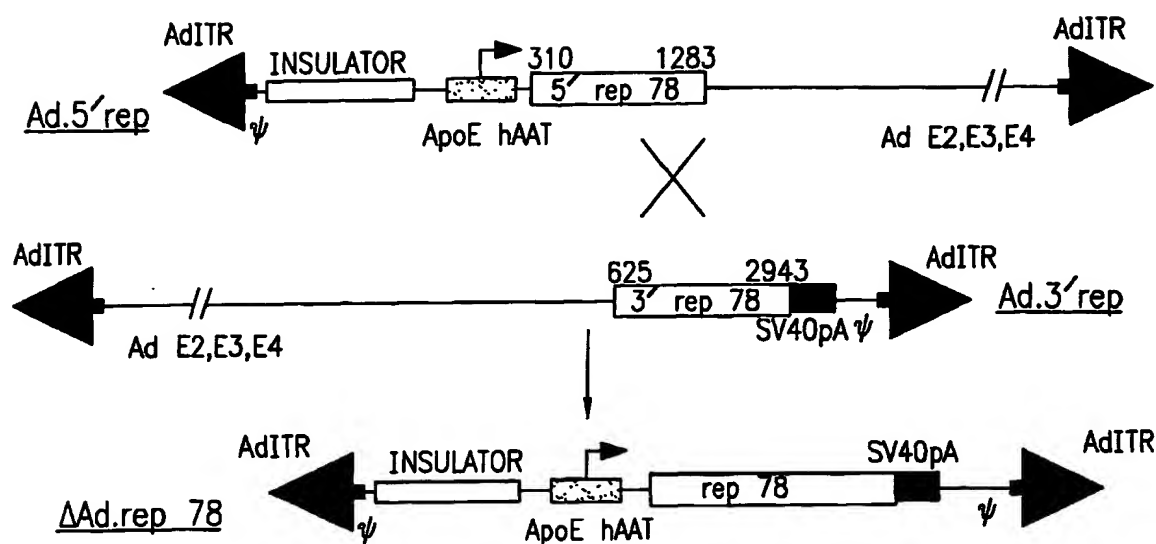


FIG. 29A

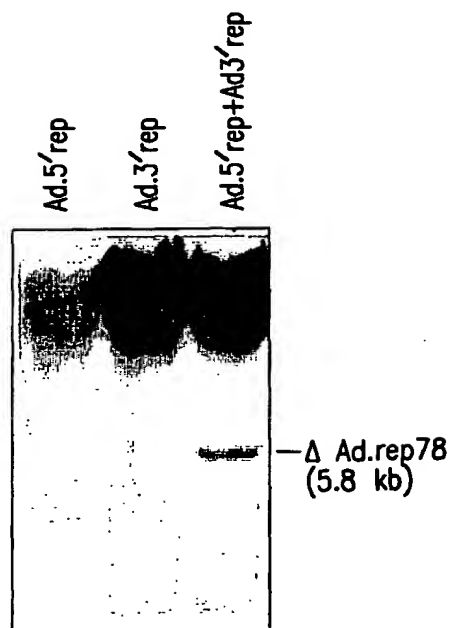


FIG. 29B

34/34

|   |   |   |   |   |   |                  |
|---|---|---|---|---|---|------------------|
| - | - | + | - | + | + | Ad-Rep5' (virus) |
| - | - | - | + | + | + | Ad-Rep3' (virus) |
| - | + | - | - | - | - | pRep78 (plasmid) |
| + | + | - | - | - | - | Ad-helper        |

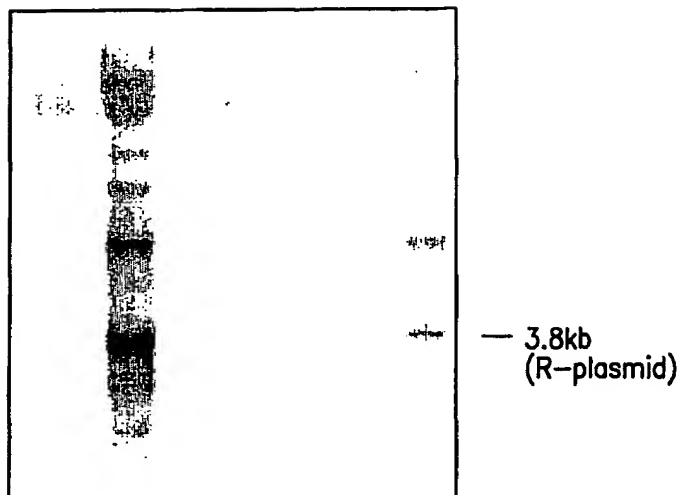


FIG. 29C

|   |   |   |   |   |                  |
|---|---|---|---|---|------------------|
| - | - | + | - | + | Ad-Rep5' (virus) |
| - | - | - | + | + | Ad-Rep3' (virus) |
| - | + | - | - | - | pRep78 (plasmid) |

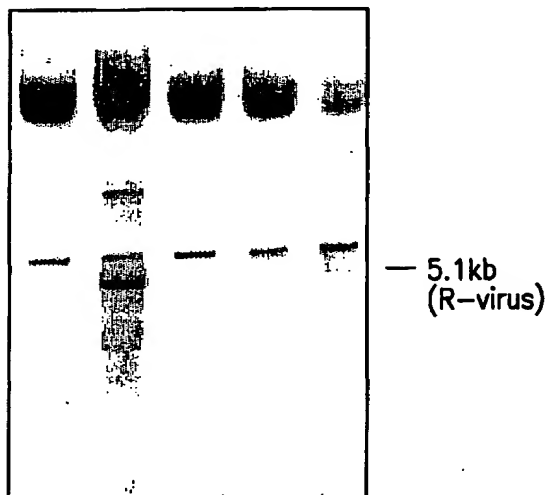


FIG. 29D

# INTERNATIONAL SEARCH REPORT

Intern 1st Application No  
PCT/US 00/15442

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12N15/861 C12N15/864 C12N15/10 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                  |
|------------|---|--|
| X          | US 5 856 152 A (WILSON JAMES M ET AL)<br>5 January 1999 (1999-01-05)  | 1,2,<br>5-11,15,<br>36-44,<br>48-54,58 |
| Y          | the whole document  | 3,4,13,<br>16-32,<br>35,36,<br>46,47   |
| Y          | WO 98 54346 A (ROELVINK PETRUS W ;EINFELD<br>DAVID (US); GENVEC INC (US); KOVESDI IM)<br>3 December 1998 (1998-12-03)<br>page 7, line 27 -page 14, line 26<br>the whole document<br>---<br>-/-- | 3,4,<br>16-32,<br>35,56                |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*A\* document member of the same patent family

Date of the actual completion of the international search

16 November 2000

Date of mailing of the international search report

30/11/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Mandl, B

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/15442

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |   |  |
|--|---|--|
| Category *   | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                  |
| X  | BURCIN M. M. ET AL.: "Adenovirus-mediated regulable target gene expression in vivo" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 96, January 1999 (1999-01), pages 355-360, XP002130326<br>ISSN: 0027-8424  | 1,2                                    |
| Y  | the whole document  | 13,46                                  |
| X  | RECCHIA A. ET AL.: "Site-specific integration mediated by a hybrid adenovirus/adeno-associated virus vector." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 96, no. 6, 16 March 1999 (1999-03-16), pages 2615-2620, XP002152917<br>ISSN: 0027-8424 | 1,2,<br>5-11,14,<br>36-44,<br>47-49,58 |
|  | the whole document  |  |
| X  | WO 97 38723 A (BARBER JACK ;IMMUSOL INC (US); LI QI XIANG (US); YU GANG (US); YU) 23 October 1997 (1997-10-23)  | 55,57                                  |
| Y  | page 4, line 14 - line 20<br>page 5, line 21 -page 6, line 10<br>page 29, line 4 -page 30, line 2<br>page 35, line 12 - line 18   | 56                                     |
| X  | WO 94 06920 A (MEDICAL RES COUNCIL ;RUSSELL STEPHEN JAMES (GB); HAWKINS ROBERT ED) 31 March 1994 (1994-03-31)   | 55,57                                  |
|  | page 19, last paragraph   |  |
| A  | GALL J. ET AL.: "ADENOVIRUS TYPE 5 AND 7 CAPSID CHIMERA: FIBER REPLACEMENT ALTERS RECEPTOR TROPISM WITHOUT AFFECTING PRIMARY IMMUNE NEUTRALIZATION EPITOPES" JOURNAL OF VIROLOGY, vol. 70, no. 4, 1 April 1996 (1996-04-01), pages 2116-2123, XP000653519<br>ISSN: 0022-538X    | 16-25                                  |
|  | the whole document  |  |
| A  | LYU Y. L. ET ALW: "Inversion/dimerization of plasmids mediated by inverted repeats." JOURNAL OF MOLECULAR BIOLOGY, vol. 285, no. 4, 29 January 1999 (1999-01-29), pages 1485-1501, XP002152918<br>ISSN: 0022-2836   | 1-58                                   |
|  | cited in the application<br>the whole document  |  |
|  | -/-   |  |

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/15442

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
|------------|--|-----------------------|
| P,Y        | STEINWAERDER D. S. ET AL.: "Generation of adenovirus vectors devoid of all viral genes by recombination between inverted repeats."<br>JOURNAL OF VIROLOGY,<br>vol. 73, no. 11, November 1999 (1999-11),<br>pages 9303-9313, XP002152919<br>ISSN: 0022-538X<br>the whole document | 1-58                  |
| P,Y        | LIEBER A. ET AL.: "Integrating adenovirus-adeno-associated virus hybrid vectors devoid of all viral genes."<br>JOURNAL OF VIROLOGY,<br>vol. 73, no. 11, November 1999 (1999-11),<br>pages 9314-9324, XP002152920<br>ISSN: 0022-538X<br>the whole document                        | 1-58                  |



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/15442

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---------------------|----------------------------|---------------------|
| US 5856152 A                              | 05-01-1999          | AU 695811 B                | 20-08-1998          |
|   |                     | AU 4405596 A               | 23-05-1996          |
|   |                     | CA 2203808 A               | 09-05-1996          |
|   |                     | EP 1046711 A               | 25-10-2000          |
|   |                     | EP 0797678 A               | 01-10-1997          |
|   |                     | JP 10507928 T              | 04-08-1998          |
|   |                     | WO 9613598 A               | 09-05-1996          |
|   |                     | US 5871982 A               | 16-02-1999          |
| WO 9854346 A                              | 03-12-1998          | AU 7604998 A               | 30-12-1998          |
|   |                     | BR 9809173 A               | 01-08-2000          |
|   |                     | EP 0988390 A               | 29-03-2000          |
|   |                     | PL 337130 A                | 31-07-2000          |
| WO 9738723 A                              | 23-10-1997          | AU 2678097 A               | 07-11-1997          |
|   |                     | CA 2251738 A               | 23-10-1997          |
|   |                     | EP 0927044 A               | 07-07-1999          |
| WO 9406920 A                              | 31-03-1994          | AU 689755 B                | 09-04-1998          |
|   |                     | AU 4827893 A               | 12-04-1994          |
|   |                     | CA 2145063 A               | 31-03-1994          |
|   |                     | EP 0670905 A               | 13-09-1995          |
|   |                     | JP 8504091 T               | 07-05-1996          |
|   |                     | US 5723287 A               | 03-03-1998          |

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number  
**WO 01/83729 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/00

(21) International Application Number: PCT/EP01/04863

(22) International Filing Date: 30 April 2001 (30.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/562,934 1 May 2000 (01.05.2000) US

(71) Applicants: NOVARTIS AG [CH/CH]; Schwarzwal-  
dallee 215, CH-4058 Basel (CH). THE SCRIPPS  
RESEARCH INSTITUTE [US/US]; 10550 North Torrey  
Pines Road, La Jolla, CA 92037 (US).

(71) Applicants and  
(72) Inventors: NEMEROW, Glen, R. [US/US]; 462 Cerro  
Street, Encinitas, CA 92024 (US). VON SEGGERN,  
Daniel, J. [US/US]; Apartment 30, 5175 Luigi Terrace,  
San Diego, CA 92122 (US). FRIEDLANDER, Marty  
[US/US]; 1720 Zapo Street, Del Mar, CA 92014 (US).

(74) Agent: BECKER, Konrad; Novartis AG, Corporate In-  
tellectual Property, Patent & Trademark Dept., CH-4002  
Basel (CH).

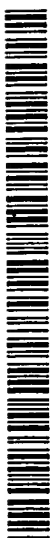
(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,  
TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished  
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.



WO 01/83729 A2

(54) Title: VECTORS FOR OCULAR TRANSDUCTION AND USE THEREOF FOR GENETIC THERAPY

(57) Abstract: Adenovirus vector-based gene therapy methods for treating ocular disorders are provided. Adenovirus vectors for therapy of ocular diseases and methods of treatment using the vectors are provided. Compositions, kits, and methods of preparation and use of the vectors for gene therapy are provided.

## **VECTORS FOR OCULAR TRANSDUCTION AND USE THEREOF FOR GENETIC THERAPY**

Work described herein was supported by NIH grants EY11431 and HL54352. The government has certain rights in such subject matter.

### **5 RELATED APPLICATIONS**

This application claims the benefit of priority to U.S. application Serial No. 09/562,934, filed May 1, 2000, to Glen R. Nemerow, Daniel Von Seggern,; Martin Friedlander, entitled "VECTORS FOR OCULAR TRANSDUCTION AND USE THEREFOR FOR GENETIC THERAPY".

- 10 This application is related to copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265, filed January 14, 2000)), to Daniel Von Seggern, Glen R. Nemerow, Paul Hallenbeck, Susan Stevenson, Yelena Skripchenko, filed January 14, 2000, entitled "Adenovirus Vectors, Packaging Cell Lines, Compositions, and
- 15 Methods for Preparation and Use," which is a continuation-in-part of U.S. Application 09/423,783 filed November 12, 1999 and claims the benefit of the filing date of U.S. Provisional Application 60/115,920 filed January 14, 1999. Where permitted, the contents and subject matter of each application and of the provisional application are incorporated in their entirety herein by reference.

### **20 FIELD OF INVENTION**

- The present invention relates to gene therapy, especially to adenovirus vector-based gene therapy. In particular, adenovirus vectors for therapy of ocular diseases and methods of treatment using the vectors are provided. Compositions, kits, and methods of preparation and use of the vectors for gene
- 25 therapy are provided.

### **BACKGROUND OF THE INVENTION**

#### **Retinal dystrophies**

- The eye is susceptible to a number of hereditary and/or age related degenerative disorders. In the United States, common causes of irreversible
- 30 blindness or severe loss of vision are retinal dystrophies (see, *e.g.*, Cotlier *et al.* (1995) *Surv. Ophthalmology* 40:51-61; Bird (1995) *Am. J. Ophthalm.* 119: 543-562; and Adler (1996) *Arch Ophthalm* 114:79-83). The retina is the sensory

-2-

tunic of the eye, containing light sensitive receptors, a complex of neurons, and pigmented epithelium, arranged in discrete layers. In humans, the macula is the portion of the retina that lies directly behind the lens. Cones, the photoreceptor cells responsible for central vision, are heavily concentrated in the macula.

5 Central dystrophies, which affect the macula, include Best's disease, age-related macular degeneration, and Stargardt's macular dystrophy. The peripheral retina is composed mainly of rods, which are responsible for side and night vision. Peripheral degenerative retinal diseases include retinitis pigmentosa, choroidemia and Bietti's crystalline dystrophy.

10 Macular degenerations are a heterogeneous group of diseases, characterized by progressive central vision loss and degeneration of the macula and underlying retinal pigmented epithelium. Age-related macular degeneration (ARMD) is the most common form of the disease, affecting an estimated 20% of persons over 75 years of age. ARMD is poorly understood in terms of etiology  
15 and pathogenesis. The very late onset of the disease has made genetic mapping particularly difficult. Certain macular degenerative conditions with a clear genetic basis, such as Stargardt's and Best's diseases, share many features with ARMD, but have been more amenable to molecular and genetic analysis.

Hereditary peripheral retinopathies are also relatively common. Retinitis  
20 pigmentosa (RP), for example, affects approximately 1.5 million people worldwide. Substantial genetic heterogeneity has been observed in this condition, with over 20 chromosomal loci identified. A predisposition to retinitis pigmentosa can be inherited by autosomal dominant, autosomal recessive, X-linked or digenic mode. Mutations have been identified in seven genes, four of  
25 which encode proteins in the rod phototransduction cascade: rhodopsin, alpha and beta subunits of rod cGMP phosphodiesterase, and rod cGMP cation-gated channel protein .alpha. subunit. Mutations in the peripherin/RDS gene have been linked to retinitis pigmentosa and macular degeneration. A single peripherin/RDS mutation apparently caused retinitis pigmentosa, pattern dystrophy and fundus  
30 flavimaculatus, in different family members.

-3-

In spite of causal heterogeneity, there is significant clinical similarity among RP subtypes. Common signs and symptoms include early electroretinographic abnormalities, ophthalmoscopic findings, and protracted, contiguous expansion of the ring-like scotoma toward the macula, leading to progressively worsening tunnel vision. A recent hypothesis is that active photoreceptor cell death, which is characteristic of these genetically distinct disorders, is mediated by a common induction of apoptosis. It may be possible to treat these conditions by the administration of agents that block induction of apoptosis in photoreceptors, such as neurotrophic factors.

10 **Adenovirus delivery vectors**

Adenovirus, which is a DNA virus with a 36 kilobase (kb) genome, is very well-characterized and its genetics and genetic organization are understood. The genetic organization of adenoviruses permits substitution of large fragments of viral DNA with foreign DNA. In addition, recombinant adenoviruses are structurally stable and no rearranged viruses are observed after extensive amplification.

Adenoviruses have been employed as delivery vehicles for introducing desired genes into eukaryotic cells. The adenovirus delivers such genes to eukaryotic cells by binding to cellular receptors followed by internalization. The adenovirus fiber protein is responsible for binding to cells. The fiber protein has two domains, a rod-like shaft portion and a globular head portion that contains the receptor binding region. The fiber spike is a homotrimer, and there are 12 spikes per virion. Human adenoviruses bind to and infect a broad range of cultured cell lines and primary tissues from different species.

25 The 35,000+ base pair (bp) genome of adenovirus type 2 has been sequenced and the predicted amino acid sequences of the major coat proteins (hexon, fiber and penton base) have been described (see, *e.g.*, Neumann *et al.*, *Gene* 69: 153-157 (1988); Herisse *et al.*, *Nuc. Acids Res.* 9: 4023-4041 (1981); Roberts *et al.*, *J. Biol. Chem.* 259: 13968-13975 (1984); Kinloch *et al.*, *J. Biol. Chem.* 259: 6431-6436 (1984); and Chroboczek *et al.*, *Viol.* 161: 549-554, 1987).

-4-

The 35,935 bp sequence of Ad5 DNA is also known and portions of many other adenovirus genomes have been sequenced. The upper packaging limit for adenovirus virions is about 105% of the wild-type genome length (see, *e.g.*, Bett, *et al.*, *J. Virol.* 67(10): 5911-21, 1993). Thus, for Ad2 and Ad5, this

5 would be an upper packaging limit of about 38kb of DNA.

Adenovirus DNA also includes inverted terminal repeat sequences (ITRs) ranging in size from about 100 to 150 bp, depending on the serotype. The inverted repeats permit single strands of viral DNA to circularize by base-pairing of their terminal sequences to form base-paired "panhandle" structures that are

10 required for replication of the viral DNA.

For efficient packaging, the ITRs and the packaging signal (a few hundred bp in length) comprise the "minimum requirement" for replication and packaging of a genomic nucleic acid into an adenovirus particle. Helper-dependent vectors lacking all viral ORFs but including these essential *cis* elements (the ITRs and

15 contiguous packaging sequence) have been constructed.

Ad vectors have several distinct advantages as gene delivery vehicles. For example, recombination of such vectors is rare; there are no known associations of human malignancies with adenoviral infections despite common human infection with adenoviruses; the genome may be manipulated to

20 accommodate foreign genes of a fairly substantial size; and host proliferation is not required for expression of adenoviral proteins. Adenovirus (Ad)-based gene delivery vectors efficiently infect many different cells and tissues. This broad tropism, however, means that gene delivery cannot be directed to a specific target cell. A large fraction of intravenously administered adenovirus is

25 retained by the liver, which could lead to undesirable side-effects. Adenovirus may potentiate immune responses. For example, Adenovirus type 5 (Ad5) also transduces dendritic cells, which present antigens very efficiently, thereby possibly exacerbating the immune response against the vector. It has been proposed that vectors with different targeting efficiencies might eliminate these

30 problems, permitting a lower total particle dose and more specific targeting (see, *e.g.*, U.S. application Serial No. 09/482,682).

-5-

The wealth of information on adenovirus structure and mechanism of infection, its efficient infection of nondividing cells, and its large genetic capacity make adenovirus a popular gene therapy vector. The wide expression of receptors to which adenovirus binds makes targeting adenovirus vectors  
5 difficult.

Hence there is a need to improve delivery and targeting of adenoviral vectors and also to provide treatments for ocular disorders. Therefore, it is an object herein to provide adenoviral vectors that specifically or selectively target cells in the eye. It is also an object herein to provide these vectors for treatment  
10 of ocular disorders.

#### SUMMARY OF THE INVENTION

Degenerative ocular diseases, such as, but not limited to, retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularization, and others (see, *e.g.*, Table below), have a genetic basis. Genes expressed in  
15 the photoreceptor cells at the back of the retina are implicated in these diseases. Provided herein are recombinant viral vectors for targeting therapeutic products to these cells.

Recombinant adenoviral vectors that include nucleic acid that permits specific binding to these photoreceptors are provided. In particular, the vector  
20 particles contain a fiber protein of Ad37 or a modified form thereof. As shown herein, fiber protein from Ad37 permits efficient infection of photoreceptor cells. Fiber proteins from other adenovirus D serotypes may also be used. In addition, the portions of the fiber protein, particularly those that interact with other viral structural proteins, such as penton, may be modified to resemble the viral source  
25 of the other structural proteins. As exemplified herein, the recombinant virus provided herein include Ad5 structural components. The N-terminus of the Ad37 fiber protein, which interacts with the penton protein, is modified to resemble the Ad5 fiber protein N-terminus to ensure production of viral particles.

The recombinant adenoviral vectors are intended for gene therapy of  
30 diseases in which genes expressed in the photoreceptors are implicated. Such diseases include, but are not limited to, degenerative ocular diseases, such as retinitis pigmentosa and Stargardt's disease. These vectors are also useful for

-6-

targeting to other ocular cells, such as conjunctival cells, which also bear receptors to which fiber from Ad37 and related serotypes bind.

The vectors will deliver therapeutic agents to the targeted cells for treatment of a variety of disorders (see *e.g.*, Tables 3 and 4, below)). The  
5 therapeutic agents are intended for expression in the photoreceptors and for secretion from the photoreceptor cells, which are surrounded on one side by choroidal vasculature, and on the other side by retinal vasculature, thereby providing a means for delivery of products. In addition, expression of growth factors, such as VEGF and others, can be used to enhance blood flow to the  
10 retina and prevent or slow the degeneration.

Therapeutic agents encoded by the recombinant adenoviral vectors include, but are not limited to, nucleic acid molecules encoding genes that are defective in certain hereditary disorders, nucleic acid molecules that encode antiangiogenics and antitumor agents for treatment of retinal  
15 disorders, such as retinoblastomas; nucleic acid molecules encoding trophic factors, such as glial cell line-derived neurotrophic factor (GDNF) and ciliary neurotrophic factor (CNTF), growth factors and growth factor inhibitors, antiapoptotic factors, such as Bcl-2 (CNTF), antitumor agents, anti-angiogenics, and genes or portions thereof for gene replacement or repair of defective genes.  
20 Hence, methods for treatment of inherited and acquired retinal diseases, including diseases involving neovascular and vascular degeneration are provided.

Methods for treating diseases involving genes expressed in photoreceptor cells are provided herein. The methods provided herein are practiced by  
25 administration of the recombinant viral vectors by any means suitable for delivery to the photoreceptors. A preferred mode of administration is intraocular injection including intravitreal and subretinal injection. Other modes of administration include, but are not limited to, intrascleral, periorbital and intravenous administration. The vectors also can include photoreceptor-specific  
30 promoters thereby providing a means, not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression.



-7-

**DETAILED DESCRIPTION OF THE INVENTION****A. DEFINITIONS**

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents, applications, published applications and other publications and sequences from GenBank and other data bases referred to anywhere in the disclosure herein are incorporated by reference in their entirety.

As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used routinely in the art (see, Table 1).

As used herein, amino acid residue refers to an amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH<sub>2</sub> refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxyl terminus of a polypeptide. In keeping with standard polypeptide nomenclature described in *J. Biol. Chem.*, 243:3552-59 (1969) and adopted at 37 C.F.R. § § 1.821 - 1.822, abbreviations for amino acid residues are shown in the following Table:

**Table 1**  
**Table of Correspondence**

| SYMBOL   |          | AMINO ACID    |
|----------|----------|---------------|
| 1-Letter | 3-Letter |               |
| Y        | Tyr      | tyrosine      |
| G        | Gly      | glycine       |
| F        | Phe      | phenylalanine |
| M        | Met      | methionine    |

-8-

5

10

15

20

| SYMBOL |     |                  |
|--------|-----|------------------|
| A      | Ala | alanine          |
| S      | Ser | serine           |
| I      | Ile | isoleucine       |
| L      | Leu | leucine          |
| T      | Thr | threonine        |
| V      | Val | valine           |
| P      | Pro | proline          |
| K      | Lys | lysine           |
| H      | His | histidine        |
| Q      | Gln | glutamine        |
| E      | Glu | glutamic acid    |
| Z      | Glx | Glu and/or Gln   |
| W      | Trp | tryptophan       |
| R      | Arg | arginine         |
| D      | Asp | aspartic acid    |
| N      | Asn | asparagine       |
| B      | Asx | Asn and/or Asp   |
| C      | Cys | cysteine         |
| X      | Xaa | Unknown or other |

It should be noted that all amino acid residue sequences represented herein by formulae have a left to right orientation in the conventional direction of amino-terminus to carboxyl-terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those referred to in 37 C.F.R. § § 1.821-1.822, and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or

-9-

more amino acid residues or to an amino-terminal group such as  $\text{NH}_2$  or to a carboxyl-terminal group such as  $\text{COOH}$ .

In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al., *Molecular Biology of the Gene*, 4th Edition, 1987, The Bejacmin/Cummings Pub. co., p.224).

Such substitutions are preferably made in accordance with those set forth in TABLE 2 as follows:

| TABLE 2 |                           |
|---------|---------------------------|
|         | Original residue          |
|         | Ala (A)                   |
| 15      | Arg (R)                   |
|         | Asn (N)                   |
|         | Cys (C)                   |
|         | Gln (Q)                   |
|         | Glu (E)                   |
| 20      | Gly (G)                   |
|         | His (H)                   |
|         | Ile (I)                   |
|         | Leu (L)                   |
|         | Lys (K)                   |
| 25      | Met (M)                   |
|         | Phe (F)                   |
|         | Ser (S)                   |
|         | Thr (T)                   |
|         | Trp (W)                   |
| 30      | Tyr (Y)                   |
|         | Val (V)                   |
|         | Conservative substitution |
|         | Gly; Ser                  |
|         | Lys                       |
|         | Gln; His                  |
|         | Ser                       |
|         | Asn                       |
|         | Asp                       |
|         | Ala; Pro                  |
|         | Asn; Gln                  |
|         | Leu; Val                  |
|         | Ile; Val                  |
|         | Arg; Gln; Glu             |
|         | Leu; Tyr; Ile             |
|         | Met; Leu; Tyr             |
|         | Thr                       |
|         | Ser                       |
|         | Tyr                       |
|         | Trp; Phe                  |
|         | Ile; Leu                  |

Other substitutions are also permissible and may be determined empirically or in accord with known conservative substitutions.

As used herein, a complementing plasmid describes plasmid vectors that deliver nucleic acids into a packaging cell line for stable integration into a chromosome in the cellular genome.

As used herein, a delivery plasmid is a plasmid vector that carries or delivers nucleic acids encoding a therapeutic gene or gene that encodes a

-10-

therapeutic product or a precursor thereof or a regulatory gene or other factor that results in a therapeutic effect when delivered *in vivo* in or into a cell line, such as, but not limited to a packaging cell line, to propagate therapeutic viral vectors.

5           As used herein, a variety of vectors with different requirements are described. For example, one vector is used to deliver particular nucleic acid molecules into a packaging cell line for stable integration into a chromosome. These types of vectors are generally identified herein as complementing plasmids. A further type of vector described herein carries or delivers nucleic acid molecules in or into a cell line (e.g., a packaging cell line) for the purpose of propagating therapeutic viral vectors; hence, these vectors are generally referred to herein as delivery plasmids. A third "type" of vector described herein is used to carry nucleic acid molecules encoding therapeutic proteins or polypeptides or regulatory proteins or are regulatory sequences to specific cells or cell types in a subject in need of treatment; these vectors are generally identified herein as therapeutic viral vectors or recombinant adenoviral vectors or viral Ad-derived vectors and are in the form of a virus particle encapsulating a viral nucleic acid containing an expression cassette for expressing the therapeutic gene.

20           As used herein, a DNA or nucleic acid homolog refers to a nucleic acid that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" is meant having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a lesser percentage of homology or identity and conserved biological activity or function.

25           The terms "homology" and "identity" are often used interchangeably. In this regard, percent homology or identity may be determined, for example, by comparing sequence information using a GAP computer program. The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443 (1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482 (1981)). Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default

-11-

parameters for the GAP program may include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745 (1986), as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN SEQUENCE*  
5 *AND STRUCTURE*, National Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

Whether any two nucleic acid molecules have nucleotide sequences that  
10 are at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% "identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988). Alternatively the BLAST function of the National Center for Biotechnology Information database may be used to  
15 determine identity.

In general, sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics*  
20 *and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).  
25 While there exist a number of methods to measure identity between two polynucleotides or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988)). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge  
30 Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988). Methods to determine identity and similarity are codified in computer programs. Preferred

-12-

computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., *et al.*, *Nucleic Acids Research* 12(II):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., *et al.*, *J Molec Biol* 215:403 (1990)).

5           Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide. For example, a test polypeptide may be defined as any polypeptide that is 90% or more identical to a reference polypeptide. As used herein, the term at least "90% identical to" refers to percent identities from 90 to 99.99 relative to the  
10   reference polypeptides. Identity at a level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared. No more than 10% (i.e., 10 out of 100) amino acids in the test polypeptide differs from that of the reference polypeptides. Similar comparisons may be made between a test and  
15   reference polynucleotides. Such differences may be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they may be clustered in one or more locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference (approximately 90% identity). Differences are defined as nucleic acid or amino acid substitutions, or  
20   deletions.

          As used herein, genetic therapy involves the transfer of heterologous DNA to the certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The DNA is introduced into the selected target cells in a manner such that the heterologous DNA is  
25   expressed and a therapeutic product encoded thereby is produced. Alternatively, the heterologous DNA may in some manner mediate expression of DNA that encodes the therapeutic product, it may encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic  
30   acid encoding a gene product to replace a defective gene or supplement a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic compound, such as a growth

-13-

factor inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous DNA encoding the therapeutic product may be modified  
5 prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof.

As used herein, heterologous DNA is DNA that encodes RNA and proteins that are not normally produced *in vivo* by the cell in which it is expressed or that mediates or encodes mediators that alter expression of endogenous DNA by  
10 affecting transcription, translation, or other regulatable biochemical processes. Heterologous DNA may also be referred to as foreign DNA. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by heterologous DNA. Examples of heterologous DNA include, but are not limited to, DNA that encodes  
15 traceable marker proteins, such as a protein that confers drug resistance, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and DNA that encodes other types of proteins, such as antibodies. Antibodies that are encoded by heterologous DNA may be secreted or expressed on the surface of the cell in which the heterologous DNA has been  
20 introduced.

Hence, herein heterologous DNA or foreign DNA, refers to a DNA molecule not present in the exact orientation and position as the counterpart DNA molecule found in the corresponding wild-type adenovirus. It may also refer to a DNA molecule from another organism or species (*i.e.*, exogenous) or  
25 from another Ad serotype.

As used herein, a therapeutically effective product is a product that is encoded by heterologous DNA that, upon introduction of the DNA into a host, a product is expressed that effectively ameliorates or eliminates the symptoms, manifestations of an inherited or acquired disease or that cures said disease.

30 Typically, DNA encoding the desired heterologous DNA is cloned into a plasmid vector and introduced by routine methods, such as calcium-phosphate mediated DNA uptake (see, (1981) Somat. Cell. Mol. Genet. 7:603-616) or

-14-

microinjection, into producer cells, such as packaging cells. After amplification in producer cells, the vectors that contain the heterologous DNA are introduced into selected target cells.

As used herein, an expression or delivery vector refers to any plasmid or virus into which a foreign or heterologous DNA may be inserted for expression in a suitable host cell — *i.e.*, the protein or polypeptide encoded by the DNA is synthesized in the host cell's system. Vectors capable of directing the expression of DNA segments (genes) encoding one or more proteins are referred to herein as "expression vectors." Also included are vectors that allow cloning of cDNA (complementary DNA) from mRNAs produced using reverse transcriptase.

As used herein, a gene is a nucleic acid molecule whose nucleotide sequence encodes RNA or polypeptide. A gene can be either RNA or DNA. Genes may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

As used herein, tropism with reference to an adenovirus refers to the selective infectivity or binding that is conferred on the particle by the fiber protein, such as by the C-terminus portion that comprises the knob.

As used herein, isolated with reference to a nucleic acid molecule or polypeptide or other biomolecule means that the nucleic acid or polypeptide has separated from the genetic environment from which the polypeptide or nucleic acid were obtained. It may also mean altered from the natural state. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Thus, a polypeptide or polynucleotide produced and/or contained within a recombinant host cell is considered isolated. Also intended as an "isolated polypeptide" or an "isolated polynucleotide" are polypeptides or polynucleotides that have been purified, partially or substantially, from a recombinant host cell or from a native source. For example, a recombinantly produced version of a compound can be substantially purified by the one-step method described in



-15-

Smith and Johnson, *Gene* 67:31-40 (1988). The terms isolated and purified are sometimes used interchangeably.

Thus, by "isolated" is meant that the nucleic acid is free of the coding sequences of those genes that, in the naturally-occurring genome of the  
5 organism (if any) immediately flank the gene encoding the nucleic acid of interest. Isolated DNA may be single-stranded or double-stranded, and may be genomic DNA, cDNA, recombinant hybrid DNA, or synthetic DNA. It may be identical to a native DNA sequence, or may differ from such sequence by the deletion, addition, or substitution of one or more nucleotides.

10 Isolated or purified as it refers to preparations made from biological cells or hosts means any cell extract containing the indicated DNA or protein including a crude extract of the DNA or protein of interest. For example, in the case of a protein, a purified preparation can be obtained following an individual technique or a series of preparative or biochemical techniques and the DNA or protein of  
15 interest can be present at various degrees of purity in these preparations. The procedures may include for example, but are not limited to, ammonium sulfate fractionation, gel filtration, ion exchange chromatography, affinity chromatography, density gradient centrifugation and electrophoresis.

A preparation of DNA or protein that is "substantially pure" or "isolated"  
20 should be understood to mean a preparation free from naturally occurring materials with which such DNA or protein is normally associated in nature. "Essentially pure" should be understood to mean a "highly" purified preparation that contains at least 95% of the DNA or protein of interest.

A cell extract that contains the DNA or protein of interest should be  
25 understood to mean a homogenate preparation or cell-free preparation obtained from cells that express the protein or contain the DNA of interest. The term "cell extract" is intended to include culture media, especially spent culture media from which the cells have been removed.

As used herein, a packaging cell line is a cell line that provides a missing  
30 gene product or its equivalent.

As used herein, an adenovirus viral particle is the minimal structural or functional unit of a virus. A virus can refer to a single particle, a stock of

-16-

particles or a viral genome. The adenovirus (Ad) particle is relatively complex and may be resolved into various substructures.

As used herein, "penton" or "penton complex" are preferentially used herein to designate a complex of penton base and fiber. The term "penton" may  
5 also be used to indicate penton base, as well as penton complex. The meaning of the term "penton" alone should be clear from the context within which it is used.

As used herein, a plasmid refers to an autonomous self-replicating extrachromosomal circular nucleic acid molecule, typically DNA.

10 As used herein, a post-transcription regulatory element (PRE) is a regulatory element found in viral or cellular messenger RNA that is not spliced, i.e. intronless messages. Examples include, but are not limited to, human hepatitis virus, woodchuck hepatitis virus, the TK gene and mouse histone gene. The PRE may be placed before a polyA sequence and after a heterologous DNA  
15 sequence.

As used herein, pseudotyping describes the production of adenoviral vectors having modified capsid protein or capsid proteins from a different serotype than the serotype of the vector itself. One example, is the production of an adenovirus 5 vector particle containing an Ad37 fiber protein. This may be  
20 accomplished by producing the adenoviral vector in packaging cell lines expressing different fiber proteins.

As used herein, promoters of interest herein may be inducible or constitutive. Inducible promoters will initiate transcription only in the presence of an additional molecule; constitutive promoters do not require the presence of  
25 any additional molecule to regulate gene expression. a regulatable or inducible promoter may also be described as a promoter where the rate or extent of RNA polymerase binding and initiation is modulated by external stimuli. Such stimuli include, but are not limited to various compounds or compositions, light, heat, stress and chemical energy sources. Inducible, suppressible and repressible  
30 promoters are considered regulatable promoters. Preferred promoters herein, are promoters that are selectively expressed in ocular cells, particularly photoreceptor cells.

-17-

As used herein, receptor refers to a biologically active molecule that specifically or selectively binds to (or with) other molecules. The term "receptor protein" may be used to more specifically indicate the proteinaceous nature of a specific receptor.

- 5       As used herein, recombinant refers to any progeny formed as the result of genetic engineering. This may also be used to describe a virus formed by recombination of plasmids in a packaging cell.

- 10       As used herein, a transgene or therapeutic nucleic acid molecule includes DNA and RNA molecules encoding an RNA or polypeptide. Such molecules may be "native" or naturally-derived sequences; they may also be "non-native" or "foreign" that are naturally- or recombinantly-derived. The term "transgene," which may be used interchangeably herein with the term "therapeutic nucleic acid molecule," is often used to describe a heterologous or foreign (exogenous) gene that is carried by a viral vector and transduced into a host cell.

- 15       Therefore, therapeutic nucleotide nucleic acid molecules include antisense sequences or nucleotide sequences which may be transcribed into antisense sequences. Therapeutic nucleotide sequences (or transgenes) all include nucleic acid molecules that function to produce a desired effect in the cell or cell nucleus into which said therapeutic sequences are delivered. For example, a therapeutic  
20       nucleic acid molecule can include a sequence of nucleotides that encodes a functional protein intended for delivery into a cell which is unable to produce that functional protein.

As used herein, the vitreous of the eye refers to material that fills the chamber behind the lens of the eye (i.e., vitreous humor or vitreous body).

- 25       As used herein, a promoter region refers to the portion of DNA of a gene that controls transcription of the DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region  
30       includes sequences that modulate this recognition, binding and transcription initiation activity of the RNA polymerase. These sequences may be *cis* acting or

-18-

may be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated.

Thus, promoters are nucleic acid fragments that contain a DNA sequence that controls the expression of a gene located 3' or downstream of the promoter. The promoter is the DNA sequence to which RNA polymerase specifically binds and initiates RNA synthesis (transcription) of that gene, typically located 3' of the promoter. A promoter also includes DNA sequences that direct the initiation of transcription, including those to which RNA polymerase specifically binds. If more than one nucleic acid sequence encoding a particular polypeptide or protein is included in a therapeutic viral vector or nucleotide sequence, more than one promoter or enhancer element may be included, particularly if that would enhance efficiency of expression.

A regulatable or inducible promoter may be described as a promoter wherein the rate of RNA polymerase binding and initiation is modulated by external stimuli. (see, *e.g.*, U.S. Patent Nos. 5,750,396 and 5,998,205). Such stimuli include various compounds or compositions, light, heat, stress, chemical energy sources, and the like. Inducible, suppressible and repressible promoters are considered regulatable promoters.

Regulatable promoters may also include tissue-specific promoters. Tissue-specific promoters direct the expression of the gene to which they are operably linked to a specific cell type. Tissue-specific promoters cause the gene located 3' of it to be expressed predominantly, if not exclusively, in the specific cells where the promoter expressed its endogenous gene. Typically, it appears that if a tissue-specific promoter expresses the gene located 3' of it at all, then it is expressed appropriately in the correct cell types (see, *e.g.*, Palmiter et al. (1986) Ann. Rev. Genet. 20: 465-499).

As used herein, the phrase "operatively linked" generally means the sequences or segments have been covalently joined into one piece of DNA, whether in single or double stranded form, whereby control sequences on one segment control expression or replication or other such control of other segments. The two segments are not necessarily contiguous.

As used herein, exogenous encompasses any therapeutic composition that is administered by the therapeutic methods provided herein. Thus, exogenous may also be referred to herein as foreign, or non-native or other equivalent expression.

5 B. Ad37 fiber tropism

The adenovirus fiber protein is a major determinant of adenovirus tropism (Gall *et al.* (1996) *J. Virol.* 70:2116-2123; Stevenson *et al.* (1995) *J. Virol.* 69:2850-2857). The fiber protein extends from the capsid and mediates viral binding to the cell surface by binding to specific cell receptors (Philipson *et al.* 10 (1968) *J. Virol.* 2:1064-1075). The fiber is a trimeric protein that includes an N-terminal tail domain that interacts with the adenovirus penton base, a central shaft domain of varying length, and a C-terminal knob domain that contains the cell receptor binding site (Chroboczek *et al.* (1995) *Curr.Top.Microbiol.Immunol.* 15 199:163-200; Riurok *et al.* (1990) *J.Mol.Biol.* 215:589-596; Stevenson *et al.* (1995) *J. Virol.* 69:2850-2857). Fiber proteins of most adenovirus subgroups have been shown to bind specifically or selectively to the 46 kDa coxsackievirus-adenovirus receptor (CAR), (Bergelson *et al.* (1997) *Science* 275:1320-1323; Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915). CAR appears to be expressed in a variety of human tissues, including the lung, at various levels (Bergelson *et al.* 20 (1997) *Science* 275:1320-1323), but Ad37 binds poorly to lung epithelial cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). This suggests that the tropism of this serotype may be influenced by factors independent of CAR expression.

Structural and biochemical data also suggest that distinct receptor binding sites are located on different regions of the Ad5 and Ad37 fiber knobs.

25 Adopting the nomenclature of Xia *et al.* (Xia *et al.* (1994) *Structure* 2:1259-1270), the receptor binding site for Ad5 is located at the AB-loop on the side of the fiber knob (Bewley *et al.* (1999) *Science* 286:1579-1583; Roelvink *et al.* (1999) *Science* 286:1568-1571). It is known that a lysine residue at position 240 of the Ad37 fiber, located in the CD-loop, is important for receptor binding 30 (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). The co-crystal structure of the Ad12 knob and the N-terminal domain of CAR (Bewley *et al.* (1999) *Science* 286:1579-1583) show that the CD-loop does not contact CAR. It thus appears

-20-

that different regions of the Ad5 and Ad37 fiber knobs recognize distinct cell receptors.

A 46 kDa receptor for coxsackieviruses and adenoviruses (CAR) mediates attachment for many adenovirus serotypes. The wide distribution of CAR fails to explain why certain adenovirus serotypes (i.e. Ad37) are highly associated with severe ocular infections such as epidemic keratoconjunctivitis (EKC). Ad37 does not use CAR, but instead uses a glycoprotein that contains sialic acid as its primary receptor (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). The modest number of Ad37 binding sites per cell (Huang *et al.* (1999) *J. Virol.* 73:2798-2802) also suggests that Ad37 recognizes a specific glycoprotein as its primary receptor for binding to conjunctival cells.

Adenovirus type 37 (subgroup D) has been associated with infections of the eye and genital tract. The tropism of Ad37 derives from the binding preference of its fiber protein, which binds to a receptor located on the surface of cells including Chang C, conjunctival epithelial cell line (Huang *et al.* (1999) *J. Virology* 73:2798-2802).

A protein receptor that is preferentially expressed on conjunctival cells to which Ad37 fiber binds is shown herein. The preferential expression of the Ad37 receptor protein on conjunctival cells suggests that this receptor likely influences Ad37 tropism and should play a key role in ocular pathogenesis. It is shown herein that Ad37 uses a distinct protein receptor that is selectively expressed on conjunctival cells. It is shown that Ad37 binds well to conjunctival cells (Chang C), but poorly to lung carcinoma cells (A549). To determine if infection correlated with cell binding, an Ad5 vector containing the Ad37 fiber protein was constructed. The 'pseudotyped' vector delivered transgenes to Chang C cells better than to A549 cells. Ad37 binding was abolished by protease treatment of Chang C cells, indicating the receptor is a membrane protein. Ad37 binding to conjunctival cells is shown herein to be calcium-dependent. It is also shown that Ad37 infection was not inhibited by a function-blocking anti-CAR monoclonal antibody, which is a feature distinct from Ad5 fiber interaction with CAR. Using a virus overlay protein blot assay (VOPBA), calcium-dependent Ad37 binding to a 50 KDa membrane protein on Chang C

-21-

cells, but not A549 cells was detected. Ad19p a closely related serotype that fails to bind to conjunctival cells, does not recognize the 50 kDa protein. Together, these data indicate that the 50 kDa protein is a candidate receptor for Ad37 on conjunctival cells.

5           Significantly, it is also shown herein that, upon administration of the vector to the vitreous humor, the recombinant adenovirus with the Ad37 fiber preferentially and selectively binds to photoreceptor cells. Hence, a recombinant adenoviral delivery vehicle that has an Ad37 fiber protein can serve as a vector for delivery of therapeutic agents to the eye for treatment of ocular disorders,  
10 including genetic and acquired disorders. The identification of the receptor for Ad37 and the resulting recognition of Ad37 tropism allows targeting of adenovirus vectors to specific human ocular cells.

As noted, fiber plays a crucial role in adenovirus infection by attaching the virus to a specific receptor on a cell surface. Hexon, penton and fiber  
15 capsomeres are the major components on the surface of the virion. The fiber is an elongated protein which exists as a trimer of three identical polypeptides (polypeptide IV) of 582 amino acids in length. An adenovirus fiber includes three domains: an N-terminal tail domain that interacts with penton base; a shaft composed of variable numbers of repeats of a 15-amino-acid segment that  
20 forms beta-sheet and beta-bends; and a knob at the C-terminus ("head domain") that contains the type-specific antigen and is responsible for binding to the cell surface receptor. The gene encoding the fiber protein from Ad2 has been expressed in human cells and has been shown to be correctly assembled into trimers, glycosylated and transported to the nucleus (see, *e.g.*, Hong and Engler,  
25 *Virology* 185: 758-761, 1991). Thus, alteration of the fiber in recombinant Ad vectors can lead to alteration in gene delivery.

As shown herein, alteration of fiber in recombinant Ad vectors such that the fiber is derived from Ad37 or another adenovirus serotype D, provides a means for selective delivery of a recombinant virus to particular cells in the eye,  
30 including conjunctival cells, and most significantly photoreceptors, thereby providing a means for targeted delivery to photoreceptor cells.

-22-

Photoreceptor cells are implicated in a number of hereditary and acquired retinal degenerative disorders. In addition, photoreceptor cells are located such that products produced therein can be delivered to other areas of the eye by virtue of the blood flow in the vicinity of the photoreceptor cells and also by virtue of the proximity of the photoreceptors to the retinal pigmented epithelium (RPE) and other retinal cells.

Hence it is contemplated herein that the recombinant viral vector will include a packaged recombinant adenovirus genome containing at least the minimal elements for replication and packaging; heterologous DNA encoding a desired gene product, typically a therapeutic product or plurality of products, such as several trophic factors, whose combined activity is effective for treating a disorder, such as a retinal degenerative disorder; and the resulting virion particles will include a fiber that has a sufficient portion to confer specific targeting to photoreceptor cells when the recombinant viral particles are introduced into the aqueous humor of a mammalian, preferably a human, eye, or otherwise contacted with the photoreceptor cells. The fiber may be a chimeric protein that has been modified for effective interaction with other coat structural proteins, such as penton. In addition, the fiber may be modified to include other elements that alter its tropism to permit binding to other cells as well (see, e.g., U.S. Patent Nos. 5,756,086 and 5,543,328, International PCT application No. WO 95/26412 and WO 98/44121 and Krasnykh, *et al.* (*J. Virol.* 70: 6839-46, 1996).

#### **C. Construction of the viral particles**

##### **1. Selection of viral genome and fiber protein**

Methods for preparing recombinant adenoviral vectors for gene product delivery are well known. Preferred among those are the methods exemplified herein (see EXAMPLES) and also described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265, filed January 14, 2000, which claims priority to U.S. provisional application Serial No. 60/115,920, as does U.S. application Serial No. 09/482,682)).



-23-

As noted, any desired recombinant adenovirus is contemplated for use in the methods herein as long as the viral genome is packaged in a capsid that includes at least the portion of a fiber protein that provides selective binding to photoreceptor cells. This fiber protein is preferably from an adenovirus type D serotype and is preferably an Ad37 fiber. The fiber protein should retain the knob region at the C-terminus ("head domain") from the Ad virus of subgroup D that contains the type-specific antigen and is responsible for binding to the cell surface receptor. Hence the fiber protein can be a chimeric fiber protein as long as it retains a sufficient portion of the type D serotype to specifically or selectively bind to photoreceptor cells. Generally the portion retained will be all or a portion of the knob region. The precise amount of knob region required can be determined empirically by including portions thereof and identifying the minimum residues from an Ad type D serotype, preferably Ad37, to effect selective targeting of a virion packaged with such fiber to photoreceptors in the eye upon introduction of the packaged virion into the aqueous humor.

Recombinant adenovirus containing heterologous nucleic acids that encode a desired product, such as a gene to correct a genetic defect, may be made by any methods known to those of skill in the art. The viruses must be packaged in a cell line that results in expression of fiber on the particles that specifically, electively or preferentially targets (binds and results in internalization) the viral particle to cells in the eye. The fiber protein from Ad37 and other Adenoviruses of serotype D that infect the eye effects such targeting. The resulting adenovirus particles that express such fiber is administered by intraocular injection, subretinal injection, particularly intravitreal injection, or any means that results in preferential accumulation in photoreceptor cells.

The family of Adenoviridae includes many members with at least 47 known serotypes of human adenovirus (Ad1-Ad47) (Shenk, *Virology*, Chapter 67, in Fields *et al.*, eds. Lippincott-Raven, Philadelphia, 1996,) as well as members of the genus Mastadenovirus including human, simian, bovine, equine, porcine, ovine, canine and opossum viruses and members of the Aviadenovirus genus, including bird viruses, such as CELO.

-24-

Thus it is contemplated that the methods herein can be applied to any recombinant viral vectors derived from any adenovirus species. One of skill in the art would have knowledge of the different adenoviruses (see, *e.g.*, Shenk, *Virology*, Chapter 67, in Fields *et al.*, eds. Lippincott-Raven, Philadelphia, 1996,) and can construct recombinant viruses containing portions of the genome of any such virus.

In the exemplified embodiment, viral particles with Ad37 fiber were prepared. Site-directed mutations were made to the Ad37 fiber gene to make the tail sequence more closely match that of Ad5 to facilitate Ad37 fiber binding to the Ad5 penton base. The plasmid for the expression of the Ad37 fiber protein, pDV80, contains the CMV promoter, the adenovirus type 5 tripartite leader (TPL), and the modified Ad37 fiber gene sequence. Genes of interest, such as nucleic acid encoding the  $\beta$  subunit of cGMP phosphodiesterase ( $\beta$ PDE),  $\beta$ -glucuronidase, rhodopsin, growth factors, anti-cancer agents, growth factor receptors and other anti-angiogenic agents, and anti-apoptotic agents, can be incorporated into these vectors using the methods known to those of skill in the art and exemplified herein.

Known adenovirus vectors, previously constructed for intraocular therapy (see, *e.g.*, Bennett *et al.* (1996) *Nature Medicine* 2:649-654, which provides an Ad virus encoding  $\beta$ PDE for treatment of retinitis pigmentosa; Cayouette *et al.* (1998) *Human Gene Therapy* 8:423-430, which provides an Ad vector that expresses CNTF for treatment of retinitis pigmentosa and other retinal degenerative diseases; and Li *et al.* (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:7700-7704, which provides an Ad virus vector that encodes a human  $\beta$ -glucuronidase for treatment of lysosomal storage disease caused by  $\beta$ -glucuronidase deficiency) can be modified by repackaging the recombinant genome using a packaging line that expresses an Ad37 fiber or other D serotype fiber.

For exemplification, nucleic acid encoding GFP was incorporated into these vectors as a means to visualize their localization. Other genes, such as genes that encode therapeutic products, may be included in place of or in addition to GFP.

-25-

Plasmid pDV80 was electroporated into E1-2a S8 cells and stable lines were selected. The fiber-deleted vectors Ad5. $\beta$ gal. $\Delta$ F and Ad5.GFP. $\Delta$ F were grown in cells in a resulting cell line, designated 705, to produce virions, which express the Ad37 fiber (Ad5. $\beta$ gal. $\Delta$ F/37F and Ad5.GFP. $\Delta$ F/37F) and CsCl-  
5 purified. These virions selectively transduce photoreceptor cells when injected intraocularly into the vitreous humor.

## 2. Packaging

Recombinant adenoviral vectors generally have at least a deletion in the first viral early gene region, referred to as E1, which includes the E1a and E1b  
10 regions. Deletion of the viral E1 region renders the recombinant adenovirus defective for replication and incapable of producing infectious viral particles in subsequently-infected target cells. Thus, to generate E1-deleted adenovirus genome replication and to produce virus particles requires a system of  
15 complementation which provides the missing E1 gene product. E1 complementation is typically provided by a cell line expressing E1, such as the human embryonic kidney packaging cell line, i.e. an epithelial cell line, called 293. Cell line 293 contains the E1 region of adenovirus, which provides E1  
20 gene region products to "support" the growth of E1-deleted virus in the cell line (see, e.g., Graham *et al.*, *J. Gen. Virol.* 36: 59-71, 1977). Additionally, cell lines that may be usable for production of defective adenovirus having a portion of the adenovirus E4 region have been reported (WO 96/22378).

Multiply deficient adenoviral vectors and complementing cell lines have also been described (WO 95/34671, U.S. Patent No. 5,994,106).

Copending U.S. application Serial No. 09/482,682 (also filed as  
25 International PCT application No. PCT/US00/00265, filed January 14, 2000)) provides packaging cell lines that support viral vectors with deletions of major portions of the viral genome, without the need for helper viruses and also provides cell lines and helper viruses for use with helper-dependent vectors. The packaging cell line has heterologous DNA stably integrated into the  
30 chromosomes of the cellular genome. The heterologous DNA sequence encodes one or more adenovirus regulatory and/or structural polypeptides that complement the genes deleted or mutated in the adenovirus vector genome to

-26-

be replicated and packaged. The packaging cell line express, for example, one or more adenovirus structural proteins, polypeptides, or fragments thereof, such as penton base, hexon, fiber, polypeptide IIIa, polypeptide V, polypeptide VI, polypeptide VII, polypeptide VIII, and biologically active fragments thereof. The expression can be constitutive or under the control of a regulatable promoter. These cell lines are designed for expression of recombinant adenoviruses intended for delivery of therapeutic products.

Particular packaging cell lines complement viral vectors having a deletion or mutation of a DNA sequence encoding an adenovirus structural protein, regulatory polypeptides E1A and E1B, and/or one or more of the following regulatory proteins or polypeptides: E2A, E2B, E3, E4, L4, or fragments thereof.

The packaging cell lines are produced by introducing each DNA molecule into the cells and then into the genome via a separate complementing plasmid or plurality of DNA molecules encoding the complementing proteins can be introduced via a single complementing plasmid. Of interest herein, is a variation in which the complementing plasmid includes DNA encoding adenovirus fiber protein (or a chimeric or modified variant thereof), from Ad virus of subgroup D, such as Ad 37, polypeptide or fragment thereof.

For therapeutic applications, the delivery plasmid further includes a nucleotide sequence encoding a foreign polypeptide. Exemplary delivery plasmids include, but are not limited to, pDV44, p $\Delta$ E1B $\beta$ -gal and p $\Delta$ E1sp1B. In a similar or analogous manner, therapeutic genes may be introduced.

The cell further includes a complementing plasmid encoding a fiber as contemplated herein; the plasmid or portion thereof is integrated into a chromosome(s) of the cellular genome of the cell.

In one embodiment, a composition comprises a cell containing first and second delivery plasmids wherein a first delivery plasmid comprises an adenovirus genome lacking a nucleotide sequence encoding fiber and incapable of directing the packaging of new viral particles in the absence of a second delivery plasmid, and a second delivery plasmid comprises an adenoviral genome capable of directing the packaging of new viral particles in the presence of the first delivery plasmid.

-27-

In a variation, the packaging cell line expresses fiber protein or chimeric variant thereof from an Ad virus of subgroup D, preferably Ad37, serotype or it can be any fiber protein but one that has been modified to include the portion of the Ad virus of subgroup D, such as Ad37, responsible for selective targeting to photoreceptors upon introduction into the vitreous humor of the eye of a mammal, preferably a human. The fiber protein can be further modified to include a non-native amino acid residue sequence that targets additional specific receptors. In all instances, the modification should not disrupt trimer formation or transport of fiber into the nucleus. In another variation, the non-native amino acid residue sequence alters the binding specificity of the fiber for a targeted cell type. The structural protein is fiber can include amino acid residue sequences from more than one adenovirus serotype. The nucleotide sequences encoding fiber protein or polypeptide need not be modified solely at one or both termini; fiber protein, may be modified "internally" as well as at the termini.

Additional nucleic acid fragments can encode polypeptides that are added to the fiber protein. In one variation, the non-native amino acid residue sequence is coupled to the carboxyl terminus of the fiber. In another, the non-native amino acid residue sequence further includes a linker sequence. Alternatively, the fiber protein further comprises a ligand coupled to the linker. Suitable ligands include, but are not limited to, ligands that specifically or selectively bind to a cell surface receptor and ligands that can be used to couple other proteins or nucleic acid molecules. Typically, the packaging cell lines will contain nucleic acid encoding the fiber protein or modified protein stably integrated into a chromosome or chromosomes in the cellular genome.

The packaging cell line can be derived from a procaryotic cell line or from a eukaryotic cell line. While various embodiments suggest the use of mammalian cells, and more particularly, epithelial cell lines, a variety of other, non-epithelial cell lines are used in various embodiments. Thus, while various embodiments disclose the use of a cell line selected from among the 293, A549, W162, HeLa, Vero, 211, and 211A cell lines, and any other cell lines suitable for such use are likewise contemplated herein.

### 3. Components of the nucleic acid molecule included in the particle

-28-

A recombinant viral vector or therapeutic viral vector for use in the methods herein, typically includes a nucleic acid fragment that encodes a protein or polypeptide molecule, or a biologically active fragment thereof, or other regulatory sequence, that is intended for use in therapeutic applications.

- 5       The nucleic acid molecule to be packaged in the viral particle also may include an enhancer element and/or a promoter located 3' or 5' to and controlling the expression of the therapeutic product-encoding nucleic acid molecule if the product is a protein. Further, for purposes herein, the promoter and/or other transcriptional and translational regulatory sequences controlling
- 10       expression of the product is preferably one that is expressed specifically in the targeted cells, such as the a photoreceptor-specific promoter, such as a rhodopsin gene promoter.

- The nucleic acid molecule to be packaged in viral capsid includes at least 2 different operatively linked DNA segments. The DNA can be manipulated and
- 15       amplified by PCR as described herein and by using standard techniques, such as those described in *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Sambrook *et al.*, eds., Cold Spring Harbor, New York (1989). Typically, to produce such molecule, the sequence encoding the selected polypeptide and the promoter or enhancer are operatively linked to a DNA molecule capable of autonomous
- 20       replication in a cell either *in vivo* or *in vitro*. By operatively linking the enhancer element or promoter and nucleic acid molecule to the vector, the attached segments are replicated along with the vector sequences.

- Thus, the recombinant DNA molecule (rDNA) is a hybrid DNA molecule comprising at least 2 nucleotide sequences not normally found together in
- 25       nature. In various preferred embodiments, one of the sequences is a sequence encoding an Ad-derived polypeptide, protein, or fragment thereof. The nucleic acid molecule intended to be packaged is from about 20 base pairs to about 40,000 base pairs in length, preferably about 50 bp to about 38,000 bp in length. In various embodiments, the nucleic acid molecule is of sufficient length
- 30       to encode one or more adenovirus proteins or functional polypeptide portions thereof. Since individual Ad polypeptides vary in length from about 19 amino acid residues to about 967 amino acid residues, encoding nucleic acid molecules

-29-

from about 50 bp up to about 3000 bp, depending on the number and size of individual polypeptide-encoding sequences that are "replaced" in the viral vectors by therapeutic product-encoding nucleic acid molecules.

Preferably the molecule includes an adenovirus tripartite leader (TPL)

- 5 nucleic acid sequence operatively linked to an intron containing RNA processing signals (such as for example, splice donor or splice acceptor sites) suitable for expression in the packaging cell line. Most preferably the intron contains a splice donor site and a splice acceptor site. Alternatively, the TPL nucleotide sequence may not comprise an intron. The intron includes any sequence of
- 10 nucleotides that function in the packaging cell line to provide RNA processing signals, including splicing signals. Introns have been well characterized from a large number of structural genes, and include but are not limited to a native intron 1 from adenovirus, such as Ad5's TPL intron 1; others include the SV40 VP intron; the rabbit beta-globin intron, and synthetic intron constructs (see,
- 15 *e.g.*, Petitclerc *et al.* (1995)*J. Biotechnol.*, 40:169; and Choi *et al.* (1991)*Mol. Cell. Biol.*, 11:3070).

- The nucleic acid molecule encoding the TPL includes either (a) first and second TPL exons or (b) first, second and third TPL exons, where each TPL exon in the sequence is selected from among the complete TPL exon 1, partial TPL
- 20 exon 1, complete TPL exon 2 and complete TPL exon 3. A complete exon is one which contains the complete nucleic acid sequence based on the sequence found in the wild type viral genome. Preferably the TPL exons are from Ad2, Ad3, Ad5, Ad7 and the like, however, they may come from any Ad serotype, as described herein. A preferred partial TPL exon 1 is described in the Examples.
- 25 The use of a TPL with a partial exon 1 has been reported (International PCT application No. WO 98/13499).

- The intron and the TPL exons can be operatively linked in a variety of configurations to provide a functional TPL nucleotide sequence. An intron may not be a part of the construct. For example, the intron can be positioned
- 30 between any of TPL exons 1, 2 or 3, and the exons can be in any order of first and second, or first/second/third. The intron can also be placed preceding the first TPL exon or following the last TPL exon. In a preferred embodiment,

-30-

complete TPL exon 1 is operatively linked to complete TPL exon 2 operatively linked to complete TPL exon 3. In a preferred variation, adenovirus TPL intron 1 is positioned between complete TPL exon 1 and complete TPL exon 2. It may also be possible to use analogous translational regulators from other viral systems such as rabiesvirus.

A preferred "complete" TPL nucleic acid molecule containing complete TPL exons 1, 2 and 3 with adenovirus intron 1 inserted between exons 1 and 2 has a nucleotide sequence shown in SEQ ID NO: 32. A preferred "partial" TPL nucleic acid molecule containing partial TPL exon 1 and complete TPL exons 2 and 3 in that order has a nucleotide sequence shown in SEQ ID NO: 26. The construction of these preferred TPL nucleotide sequences is described in the Examples.

Thus, preferred expression cassettes and complementing plasmids for expressing adenovirus structural genes, particularly fiber protein, contain an adenovirus TPL nucleotide sequence as described herein.

#### 4. Complementing Plasmids

Also contemplated are the use of nucleic acid molecules, typically in the form of DNA plasmid vectors, which are capable of expression of an adenovirus structural protein or regulatory protein. Because these expression plasmids are used to complement the defective genes of a recombinant adenovirus vector genome, the plasmids are referred to as complementing or complementation plasmids.

The complementing plasmid contains an expression cassette, a nucleotide sequence capable of expressing a protein product encoded by the nucleic acid molecule. Expression cassettes typically contain a promoter and a structural gene operatively linked to the promoter. The complementing plasmid can further include a sequence of nucleotides encoding TPL nucleotide to enhance expression of the structural gene product when used in the context of adenovirus genome replication and packaging.

A complementing plasmid can include a promoter operatively linked to a sequence of nucleotides encoding an adenovirus structural polypeptide, such as, but are not limited to, penton base; hexon; fiber; polypeptide IIIa; polypeptide V;



-31-

polypeptide VI; polypeptide VII; polypeptide VIII; and biologically active fragments thereof. In another variation, a complementing plasmid may also include a sequence of nucleotides encoding a first adenovirus regulatory polypeptide, a second regulatory polypeptide, and/or a third regulatory polypeptide, and any combination of the foregoing.

Plasmid pDV80 is a preferred plasmid herein. Other plasmids constructed in an analogous manner to encode modified fiber proteins and chimeric fiber proteins are also contemplated herein.

#### 5. Nucleic Acid Molecule Synthesis

A nucleic acid molecule comprising synthetic oligonucleotides can be prepared using any suitable method, such as the phosphotriester or phosphodiester methods (see, *e.g.*, Narang (1979) *et al.*, *Meth. Enzymol.*, 68:90; U.S. Patent No. 4,356,270; and Brown *et al.*, (1979) *Meth. Enzymol.*, 68:109). For oligonucleotides, the synthesis of the family members can be conducted simultaneously in a single reaction vessel, or can be synthesized independently and later admixed in preselected molar ratios. For simultaneous synthesis, the nucleotide residues that are conserved at preselected positions of the sequence of the family member can be introduced in a chemical synthesis protocol simultaneously to the variants by the addition of a single preselected nucleotide precursor to the solid phase oligonucleotide reaction admixture when that position number of the oligonucleotide is being chemically added to the growing oligonucleotide polymer. The addition of nucleotide residues to those positions in the sequence that vary can be introduced simultaneously by the addition of amounts, preferably equimolar amounts, of multiple preselected nucleotide precursors to the solid phase oligonucleotide reaction admixture during chemical synthesis. For example, where all four possible natural nucleotides (A,T,G and C) are to be added at a preselected position, their precursors are added to the oligonucleotide synthesis reaction at that step to simultaneously form four variants (see, *e.g.*, Ausubel *et al.* (*Current Protocols in Molecular Biology*, Suppl. 8. p.2.11.7, John Wiley & Sons, Inc., New York, 1991)).

-32-

Nucleotide bases other than the common four nucleotides (A,T,G or C), or the RNA equivalent nucleotide uracil (U), can also be used. For example, it is well known that inosine (I) is capable of hybridizing with A, T and G, but not C. Examples of other useful nucleotide analogs are known in the art and may be  
5 found referred to in 37 C.F.R. §1.822.

Thus, where all four common nucleotides are to occupy a single position of a family of oligonucleotides, that is, where the preselected nucleotide sequence is designed to contain oligonucleotides that can hybridize to four sequences that vary at one position, several different oligonucleotide structures  
10 are contemplated. The composition can contain four members, where a preselected position contains A,T,G or C. Alternatively, a composition can contain two nucleotide sequence members, where a preselected position contains I or C, and has the capacity to hybridize at that position to all four possible common nucleotides. Finally, other nucleotides may be included at the  
15 preselected position that have the capacity to hybridize in a non-destabilizing manner with more than one of the common nucleotides in a manner similar to inosine.

Similarly, larger nucleic acid molecules can be constructed in synthetic oligonucleotide pieces, and assembled by complementary hybridization and  
20 ligation, as is well known.

#### **D. Adenovirus Expression Vector Systems**

The adenovirus vector genome that is encapsulated in the virus particle and that expresses exogenous genes in a gene therapy setting is a key component of the system. Thus, the components of a recombinant adenovirus  
25 vector genome include the ability to express selected adenovirus structural genes, to express a desired exogenous protein, and to contain sufficient replication and packaging signals that the genome is packaged into a gene delivery vector particle. The preferred replication signal is an adenovirus inverted terminal repeat  
30 containing an adenovirus origin of replication, as is well known and described herein.

-33-

Although adenovirus include many proteins, not all adenovirus proteins are required for assembly of a recombinant adenovirus particle (vector). Thus, deletion of the appropriate genes from a recombinant Ad vector permits accommodation of even larger "foreign" DNA segments.

5        A preferred recombinant adenovirus vector genome is "helper independent" so that genome can replicate and be packaged without the help of a second, complementing helper virus. Complementation is provided by a packaging cell.

10       In a preferred embodiment, the adenovirus vector genome does not encode a functional adenovirus fiber protein. A non-functional fiber gene refers to a deletion, mutation or other modification to the adenovirus fiber gene such that the gene does not express any or insufficient adenovirus fiber protein to package a fiber-containing adenovirus particle without complementation of the fiber gene by a complementing plasmid or packaging cell line. Such a genome is  
15       referred to as a "fiberless" genome, not to be confused with a fiberless particle. Alternatively, a fiber protein may be encoded but is insufficiently expressed to result in a fiber containing particle.

20       Thus, contemplated for use are helper-independent fiberless recombinant adenovirus vector genomes that include genes that (a) express all adenovirus structural gene products but express insufficient adenovirus fiber protein to package a fiber-containing adenovirus particle without complementation of said fiber gene, (b) express an exogenous protein, and (c) contain an adenovirus packaging signal and inverted terminal repeats containing adenovirus origin of replication.

25       The adenovirus vector genome is propagated in the laboratory in the form of rDNA plasmids containing the genome, and upon introduction into an appropriate host, the viral genetic elements provide for viral genome replication and packaging rather than plasmid-based propagation. Exemplary methods for preparing an Ad-vector genome are described in the Examples.

30       A vector herein includes a nucleic acid (preferably DNA) molecule capable of autonomous replication in a cell and to which a DNA segment, *e.g.*, a gene or polynucleotide, can be operatively linked to bring about replication of the

-34-

attached segment. For purposes herein, one of the nucleotide segments to be operatively linked to vector sequences encodes at least a portion of a therapeutic nucleic acid molecule. As noted above, therapeutic nucleic acid molecules include those encoding proteins and also those that encode regulatory factors that can lead to expression or inhibition or alteration of expression of a gene product in a targeted cell.

#### 1. Nucleic Acid Gene Expression Cassettes

In various embodiments, a peptide-coding sequence of the therapeutic gene is inserted into an expression vector and expressed; however, it is also feasible to construct an expression vector which also includes some non-coding sequences as well. Preferably, however, non-coding sequences are excluded. Alternatively, a nucleotide sequence for a soluble form of a polypeptide may be utilized. Another preferred therapeutic viral vector includes a nucleotide sequence encoding at least a portion of a therapeutic nucleotide sequence operatively linked to the expression vector for expression of the coding sequence in the therapeutic nucleotide sequence.

The choice of viral vector into which a therapeutic nucleic acid molecule is operatively linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, vector replication and protein expression, and the host cell to be transformed – these being limitations inherent in the art of constructing recombinant DNA molecules. Although certain adenovirus serotypes are recited herein in the form of specific examples, it should be understood that the use of *any* adenovirus serotype, including hybrids and derivatives thereof are contemplated.

A translatable nucleotide sequence is a linear series of nucleotides that provide an uninterrupted series of at least 8 codons that encode a polypeptide in one reading frame. Preferably, the nucleotide sequence is a DNA sequence. The vector itself may be of any suitable type, such as a viral vector (RNA or DNA), naked straight-chain or circular DNA, or a vesicle or envelope containing the nucleic acid material and any polypeptides that are to be inserted into the cell.

-35-

## 2. Promoters

As noted elsewhere herein, an expression nucleic acid in an Ad-derived vector may also include a promoter, particularly a tissue or cell specific promoter, preferably one expressed in ocular cells, particularly photoreceptors.

5 Promoters contemplated for use herein include regulatable (inducible) as well as constitutive promoters, which may be used, either on separate vectors or on the same vector. Some useful regulatable promoters are those of the CREB-regulated gene family and include inhibin, gonadotropin, cytochrome c, glucagon, and the like. (See, *e.g.*, International PCT application No. WO  
10 96/14061). Preferably the promoter selected is from a photoreceptor-specific gene, such as a rhodopsin gene or gene that encodes a protein that regulates rhodopsin expression.

### E. Formulation and administration

Compositions containing therapeutically effective concentrations of  
15 recombinant adenovirus delivery vectors are provided. These are for delivery of therapeutic gene products to cells, particularly cells express a particular 50 kDa receptor or other receptor with which the vectors interact. These cells include cells of the eye and genital tract. Of particular interest are photoreceptor cells of the eye. Administration is effected by any means through which contacting with  
20 the photoreceptors is effected. Preferable modes of administration include, but are not limited to, subretinal injection, particularly intravitreal injection, to provide access to photoreceptor cells.

The recombinant viral compositions may also be formulated for  
implantation into the anterior or posterior chamber of the eye, preferably the  
25 vitreous cavity, in sustained released formulations, such as those adsorbed to biodegradable supports, including collagen sponges, or in liposomes. Sustained release formulations may be formulated for multiple dosage administration, so that during a selected period of time, such as a month or up to about a year, several dosages are administered. Thus, for example, liposomes may be  
30 prepared such that a total of about two to up to about five or more times the single dosage is administered in one injection.

-36-

The vectors are formulated in an ophthalmologically acceptable carrier for intraocular, preferably intravitreal, administration in a volume of between about 0.05 ml and 0.150 ml, preferably about 0.05 and 0.100 ml.

5 The composition can be provided in a sealed sterile vial containing an amount of a compound of formula I, that upon intraocular administration will deliver a sufficient amount of viral particles to the photoreceptors in a volume of about 50 to 150  $\mu$ l, containing at least about  $10^7$ , more preferably at least about  $10^8$  plaque forming units in such volume. Typically, the vials will, thus, contain about 0.150 ml of the composition.

10 To prepare compositions the viral particles are dialyzed into a suitable ophthalmologically acceptable carrier or viral particles, for example, may be concentrated and/or mixed therewith. The resulting mixture may be a solution, suspension or emulsion. In addition, the viral particles may be formulated as the sole pharmaceutically active ingredient in the composition or may be  
15 combined with other active agents for the particular disorder treated.

For administration by intraocular injection or via eyedrops, suitable carriers include, but are not limited to, physiological saline, phosphate buffered saline (PBS), balanced salt solution (BSS), lactate Ringers solution, and solutions containing thickening and solubilizing agents, such as glucose, polyethylene  
20 glycol, and polypropylene glycol and mixtures thereof. Liposomal suspensions may also be suitable as pharmaceutically acceptable carriers. These may be prepared according to methods known to those skilled in the art. Suitable ophthalmologically acceptable carriers are known. Solutions or mixtures intended for ophthalmic use may be formulated as 0.01% - 10% isotonic  
25 solutions, pH about 5-7, with appropriate salts [see, e.g., U.S. Patent No. 5,116,868, which describes typical compositions of ophthalmic irrigation solutions and solutions for local application]. Such solutions, which have a pH adjusted to about 7.4, contain, for example, 90-100 mM sodium chloride, 4-6 mM dibasic potassium phosphate, 4-6 mM dibasic sodium phosphate, 8-12 mM  
30 sodium citrate, 0.5-1.5 mM magnesium chloride, 1.5-2.5 mM calcium chloride, 15-25 mM sodium acetate, 10-20 mM D.L.-sodium  $\beta$ -hydroxybutyrate and 5-5.5 mM glucose.

-37-

The compositions may be prepared with carriers that protect them from rapid elimination from the body, such as time release formulations or coatings. Such carriers include controlled release formulations, such as, but not limited to, microencapsulated delivery systems, and biodegradable, biocompatible polymers, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, polyorthoesters, polylactic acid and other types of implants that may be placed directly into the anterior or posterior chamber or vitreous cavity of the eye. The compositions may also be administered in pellets, such as Elvax pellets (ethylene-vinyl acetate copolymer resin).

10       Liposomal suspensions, including tissue-targeted liposomes, may also be suitable as pharmaceutically acceptable carriers. For example, liposome formulations may be prepared by methods known to those of skill in the art [see, e.g., Kimm *et al.* (1983) *Bioch. Bioph. Acta* 728:339-398; Assil *et al.* (1987) *Arch Ophthalmol.* 105:400; and U.S. Patent No. 4,522,811]. The viral particles  
15       may be encapsulated into the aqueous phase of liposome systems.

      The active materials can also be mixed with other active materials, that do not impair the desired action, or with materials that supplement the desired action or have other action, including viscoelastic materials, such as hyaluronic acid, which is sold under the trademark HEALON, which is a solution of a high  
20       molecular weight (MW) of about 3 millions fraction of sodium hyaluronate [manufactured by Pharmacia, Inc; see, e.g., U.S. Patent Nos. 5,292,362, 5,282,851, 5,273,056, 5,229,127, 4,517,295 and 4,328,803], VISCOAT [fluorine-containing (meth)acrylates, such as, 1H,1H,2H,2H-hepta-decafluorodecylmethacrylate; see, e.g., U.S. Patent Nos. 5,278,126, 5,273,751  
25       and 5,214,080; commercially available from Alcon Surgical, Inc.], ORCOLON [see, e.g., U.S. Patent No. 5,273,056; commercially available from Optical Radiation Corporation], methylcellulose, methyl hyaluronate, polyacrylamide and polymethacrylamide [see, e.g., U.S. Patent No. 5,273,751]. The viscoelastic  
30       materials are present generally in amounts ranging from about 0.5 to 5.0%, preferably 1 to 3% by weight of the conjugate material and serve to coat and protect the treated tissues. The compositions may also include a dye, such as

-38-

methylene blue or other inert dye, so that the composition can be seen when injected into the eye. Additional active agents may be included.

The compositions can be enclosed in ampules, disposable syringes or multiple or single dose vials made of glass, plastic or other suitable material.

5 Such enclosed compositions can be provided in kits. In particular, kits containing vials, ampules or other containers, preferably disposable vials with sufficient amount of the composition to deliver about 0.100 ml thereof, and disposable needles, preferably self sealing 25-30 gauge needles, are provided herein.

10 Finally, the compounds may be packaged as articles of manufacture containing packaging material, typically a vial, an ophthalmologically acceptable composition containing the viral particles and a label that indicates the therapeutic use of the composition.

Also provided are kits for practice of the methods herein. The kits  
15 contain one or more containers, such as sealed vials, with sufficient composition for single dosage administration, and one or more needles, such as self sealing 25-33 gauge needles, preferably 33 gauge or smaller needles, precisely calibrated syringes or other precisely calibrated delivery device, suitable for intravitreal injection.

20 Administration of the composition is preferably by intraocular injection, although other modes of administration may be effective, if the sufficient amount of the compound achieves contact with the vitreous cavity. Intraocular injection may be effected by intravitreal injection, aqueous humor injection or injection into the external layers of the eye, such as subconjunctival injection or  
25 subtenon injection, or by topical application to the cornea, if a penetrating formulation is used.

#### Administration

The compositions containing the compounds are administered intraocularly or by other means, such as topically in the form of penetrating  
30 eyedrops, whereby contact of the recombinant vectors with the aqueous humor is effected. Intraocular administration may be effected by intravitreal injection, aqueous humor injection, injection into the external layers of the eye, such as



-39-

subconjunctival injection or subtenon injection, preferably in free form, but, alternatively, in liposomes or other sustained drug delivery device.

Administration is preferably by intravitreal injection, preferably through self sealing 25-30 gauge needles or other suitably calibrated delivery device.

- 5 Injection into the eye may be through the pars plana via the self-sealing needle.

It is further understood that, for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the recombinant viruses, and that the concentration ranges set

- 10 forth herein are exemplary only and are not intended to limit the scope or practice of the claimed methods

**F. Diseases, Disorders and therapeutic products**

**1. Disease and disorders**

**Retinitis pigmentosa**

- 15 Methods for specifically or selectively targeting recombinant adenovirus vectors for delivery of gene products, particularly therapeutic products are provided herein. These methods are particularly suitable for targeting cells that express receptors that are selectively recognized by Ad virus of subgroup D viruses, particularly Ad37. It is shown herein that these viruses selectively
- 20 recognize receptors on cells, such as conjunctival cells and photoreceptors, that are not recognized by other adenoviruses. Hence, methods for targeting to these cell types by providing vectors that are packaged in viral particles that contain a sufficient portion of a fiber protein from one of these Ad serotypes to bind to these receptors. These methods are useful for targeting to
- 25 photoreceptors and for treating ocular disorders, including, but are not limited to, inherited and acquired retinal, neovascular degenerative diseases (see table below).

- It is estimated that 1 in 3,500 individuals in the United States suffer from one of the pigmented retinopathies. This group of retinal diseases, commonly
- 30 called retinitis pigmentosa, is characterized by progressive loss of peripheral and night vision. Patients may be affected at almost any age and it is not uncommon to experience symptoms in early childhood in certain inherited forms.

-40-

It has been shown that there are a variety of mutations in genes expressed in the photoreceptors, including genes in the rhodopsin gene and pathway that appear to be responsible for these diseases. In addition to mutations in rhodopsin, changes in the retinal pigmented epithelial (RPE) cells, also undergo  
5 degenerative changes and can form clumps of pigment that give rise to the characteristic pigmentary changes seen in patients with RP.

#### Angiogenesis and ocular diseases and disorders

The vast majority of diseases that cause catastrophic loss of vision do so as a result of ocular neovascularization; age related macular degeneration  
10 (ARMD) affects 12-15 million American over the age of 65 and causes visual loss in 10-15% of them as a direct effect of choroidal (sub-retinal) neovascularization. The leading cause of visual loss for Americans under the age of 65 is diabetes; 16 million individuals in the United States are diabetic and 40,000 per year suffer from ocular complications of the disease, which often are  
15 a result of retinal neovascularization. Laser photocoagulation has been effective in preventing severe visual loss in subgroups of high risk diabetic patients, but the overall 10 year incidence of retinopathy remains essentially unchanged. For patients with choroidal neovascularization due to ARMD or inflammatory eye disease, such as ocular histoplasmosis, photocoagulation, with few exceptions,  
20 is ineffective in preventing visual loss. While recently developed, non-destructive photodynamic therapies hold promise for temporarily reducing individual loss in patients with previously untreatable choroidal neovascularization, only 61.4% of patients treated every 3-4 months had improved or stabilized vision compared to 45.9% of the placebo-treated group.

25 In the normal adult, angiogenesis is tightly regulated and limited to wound healing, pregnancy and uterine cycling. Angiogenesis is turned on by specific angiogenic molecules such as basic and acidic fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), angiogenin, transforming growth factor (TGF), tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) and platelet derived growth factor  
30 (PDGF). Angiogenesis can be suppressed by inhibitory molecules such as interferon- $\alpha$ , thrombospondin-1, angiostatin and endostatin. It is the balance of these naturally occurring stimulators and inhibitors that controls the normally

-41-

quiescent capillary vasculature. When this balance is upset, as in certain disease states, capillary endothelial cells are induced to proliferate, migrate and ultimately differentiate.

Angiogenesis plays a central role in a variety of diseases, including, but  
5 are not limited to, cancer and ocular neovascularization. Sustained growth and metastasis of a variety of tumors has also been shown to be dependent on the growth of new host blood vessels into the tumor in response to tumor derived angiogenic factors. Proliferation of new blood vessels in response to a variety of stimuli occurs as the dominant finding in the majority of eye diseases that blind,  
10 such as, but are not limited to, proliferative diabetic retinopathy (PDR), ARMD, rubeotic glaucoma, interstitial keratitis and retinopathy of prematurity. In these diseases, tissue damage can stimulate release of angiogenic factors resulting in capillary proliferation. VEGF plays a dominant role in iris neovascularization and neovascular retinopathies. While reports clearly show a correlation between  
15 intraocular VEGF levels and ischemic retinopathic ocular neovascularization, FGF likely plays a role. Basic and acidic FGF are known to be present in the normal adult retina, even though detectable levels are not consistently correlated with neovascularization. This may be largely due to the fact that FGF binds very tightly to charged components of the extracellular matrix and may not be readily  
20 available in a freely diffusible form that would be detected by standard assays of intraocular fluids.

A final common pathway in the angiogenic response involves integrin-mediated information exchange between a proliferating vascular endothelial cell and the extracellular matrix. This class of adhesion receptors, called integrins,  
25 are expressed as heterodimers having an  $\alpha$  and  $\beta$  subunit on all cells. One such integrin,  $\alpha_v\beta_3$ , is the most promiscuous member of this family and allows endothelial cells to interact with a wide variety of extracellular matrix components. Peptide and antibody antagonists of this integrin inhibit angiogenesis by selectively inducing apoptosis of the proliferating vascular  
30 endothelial cells. Two cytokine-dependent pathways of angiogenesis exist and may be defined by their dependency on distinct vascular cell integrins,  $\alpha_v\beta_3$  and  $\alpha_v\beta_5$ . Specifically, basic FGF- and VEGF-induced angiogenesis depend on integrin

-42-

$\alpha_v\beta_3$  and  $\alpha_v\beta_5$ , respectively, since antibody antagonists of each integrin selectively block one of these angiogenic pathways in the rabbit corneal and chick chorioallantoic membrane (CAM) models. Peptide antagonists that block all  $\alpha_v$  integrins inhibit FGF- and VEGF-stimulated angiogenesis. While normal human ocular blood vessels do not display either integrin,  $\alpha_v\beta_3$  and  $\alpha_v\beta_5$  integrins are selectively displayed on blood vessels in tissues from patients with active neovascular eye disease. While only  $\alpha_v\beta_3$  was consistently observed in tissue from patients with ARMD,  $\alpha_v\beta_3$  and  $\alpha_v\beta_5$  were present in tissues from patients with PDR. Systemically administered peptide antagonists of integrins blocked new blood vessel formation in a mouse model of retinal vasculogenesis.

In addition to adhesion events described above, cell migration through the extracellular matrix also depends on proteolysis. Matrix metalloproteinases are a family of zinc-requiring matrix-degrading enzymes that include the collagenases, gelatinases and stromelysins, all of which have been implicated in invasive cell behavior. Invasive cell processes such as tumor metastasis and angiogenesis have been found to be associated with the expression of integrins and MMP-2, MMP-2 are all found throughout the eye where they may interact to maintain a quiescent vasculature until the balance is upset, resulting in pathological angiogenesis. A non-catalytic C-terminal hemopexin-like domain of MMP-2 (PEX) can block cell surface collagenolytic activity and inhibit angiogenesis in the CAM model by preventing localization of MMP-2 to the surface of invasive cells through interaction with the integrin  $\alpha_v\beta_3$ .

Hence, anti-angiogenic agents have a role in treating retinal degeneration to prevent the damaging effects of these trophic and growth factors. Angiogenic agents, also have a role in promoting desirable vascularization to retard retinal degeneration by enhancing blood flow to cells.

Members of adenovirus subgroup D, Ad8, 19A, and 37, are infectious agents that cause particularly severe cases of epidemic keratoconjunctivitis (EKC) (Arnberg *et al.* (1998) *Virology* 227:239-244; Curtis *et al.* (1998) *J. Med. Microbiol.* 47:91-94; Ritterband *et al.* (1998) *Rev. Med. Virol.* 8:187-201; and Takeuchi *et al.* (1999) *J. Clin. Microbiol.* 37:3392-3394). There is no effective treatment for this debilitating and contagious disease and EKC

-43-

continues to be a problem in ophthalmology clinics worldwide (Curtis *et al.* (1998) *J.Med.Microbiol.* 47:91-94, Lukashok *et al.* (1998) *Curr.Clin.Top.Infec.Dis.* 18:286-304). Hence the vectors herein may be used for treating the disease.

5

**Table 3**  
Candidate targets for ocular disease therapy

| CANDIDATE TARGETS FOR OCULAR DISEASE THERAPY |   |
|--|---|
| Disease                                      | Candidate target(s)   |
| Retinitis pigmentosa                         | Rhodopsin gene, and genes that regulate expression thereof<br><i>rds/peripherin</i> |
| 10 Stargardt's disease                       | rim protein (ARC protein)   |
|  |   |
| Choroideremia                                | rab geranylgeranyl transferase<br>CHM, TCD, CHML *                                  |
| Gyrate Atrophy                               | ornithine aminotransferase  |
| 15 Macular dystrophy                         | <i>rds/peripherin</i>   |

15

\* see, "MSR6-yeast homologue of the choroideraemia gene," *Nature Genetics* 3: 193-4 (1993)

TABLE 4

| Other Diseases |   |
|----------------|---|
| 20             | Exudative Choroidal Diseases  |
|                | ICSC, fluorescein angiogram   |
|                | ICSC with large serous detachment of RPE (retinal pigmented epithelium) |
|                | ICSC with bullous retinal detachment                                    |
|                | Macular drusen, exudative, confluent                                    |
| 25             | Drusen, sub-RPE choroidal neovascularization                            |
|                | Drusen, notched serous detachment of RPE                                |

| Other Diseases |   |
|----------------|---|
|                | Drusen, notched serous and hemorrhagic detachment of RPE            |
|                | Drusen, serous and hemorrhagic detachment of RPE and retina         |
|                | Drusen, organized RPE detachment causing bullous retinal detachment |
|                | Drusen, geographic atrophy of RPE                                   |
| 5              | Drusen, exudative and cuticular, vitelliform macular detachment     |
|                | Drusen, cuticular, large vitelliform macular detachment             |
|                | North Carolina dystrophy with macular staphyloma                    |
|                | North Carolina dystrophy with macular staphyloma                    |
|                | Angioid streaks, pseudoxanthoma elasticum (PXE), CNVM               |
| 10             | Angioid streaks, PXE, large notched retinal detachment              |
|                | Myopic degeneration, Foerster-Fuchs spot                            |
|                | Presumed ocular histoplasmosis syndrome (POHS)                      |
|                | Submacular bacterial abscess  |
|                | <i>Toxocara canis</i> , subretinal granuloma                        |
| 15             | Serpiginous (geographic) choroiditis                                |
|                | Posterior scleritis   |
|                | Harada's disease  |
|                | Posterior sympathetic uveitis                                       |
|                | Benign reactive lymphoid hyperplasia of uveal tract                 |
| 20             | Choroidal ruptures and CNVM   |
|                | Cavernous hemangioma of choroid                                     |
|                | Choroidal osteoma   |
|                | Choroidal nevus, serous macular detachment                          |
|                | Choroidal nevus with CNVM   |
| 25             | Diffuse sclerochoroidal melanocytic nevus                           |
|                | Choroidal melanoma with serous detachment of RPE                    |
|                | Metastatic lung carcinoma to choroid                                |
|                | Sub-RPE reticulum cell sarcoma                                      |

-45-

|    |  |
|----|--|
|    | <b>Other Diseases</b>  |
|    | RPE tear, idiopathic choroidal neovascularization                              |
|    | <b>Heredodystrophic Disorders Affecting RPE &amp; Retina</b>                   |
|    | Best's vitelliform macular dystrophy   |
|    | Best's vitelliform macular dystrophy with CNVM                                 |
| 5  | Best's vitelliform macular dystrophy, multiple lesions                         |
|    | Adult-onset vitelliform foveomacular dystrophy                                 |
|    | Pattern dystrophy simulating fundus flavimaculatus                             |
|    | Stargardt's disease (fundus flavimaculatus)                                    |
|    | Asteroid macular dystrophy   |
| 10 | Sjögren-Larssen syndrome   |
|    | Oguchi's disease, light-adapted state  |
|    | Oguchi's disease, dark-adapted state   |
|    | Fundus albipunctatus   |
|    | Retinitis pigmentosa, cystoid macular edema                                    |
| 15 | Crystalline tapetoretinal dystrophy  |
|    | Choroideremia  |
|    | Goldmann-Favre syndrome  |
|    | Sex-linked juvenile retinoschisis  |
|    | Perivenous retinitis pigmentosa  |
| 20 | <b>Retinal Vascular Disorders</b>  |
|    | Retinal arteriovenous aneurysm   |
|    | Central retinal artery occlusion   |
|    | Cilioretinal artery obstruction  |
|    | Ischemic retinopathy in systemic lupus erythematosus                           |
| 25 | Ischemic retinopathy in scleroderma  |
|    | Hemorrhagic detachment of internal limiting membrane, hypertensive retinopathy |
|    | Acquired retinal arterial macroaneurysm  |

| Other Diseases                                  |  |
|---|--|
|   | Cystoid macular edema, aphakic                             |
|   | Cystoid macular edema, nicotinic acid maculopathy          |
|   | Congenital retinal telangiectasis                          |
|   | Acquired bilateral juxtafoveal telangiectasis              |
| 5   | Acquired bilateral juxtafoveal obliterative telangiectasis |
|   | Diabetic optic neuropathy                                  |
|   | X-ray radiation exudative retinopathy                      |
|   | Sickle cell SC disease, macular hemorrhage                 |
|   | Retinal arterial aneurysms, arteritis, neuroretinitis      |
| 10  | Branch retinal vein obstruction (BRVO)                     |
|   | BRVO, exudative maculopathy                                |
|   | BRVO, optic disc new vessels, photocoagulation             |
|   | Waldenström's macroglobulinemia                            |
| Inflammatory Diseases of the Retina and Choroid |  |
| 15  | Luetic retinal vasculitis                                  |
|   | Focal <i>Candida</i> retinal abscess                       |
|   | Toxoplasmosis, atrophic chorioretinal scar                 |
|   | Toxoplasmosis retinitis and macular detachment             |
|   | Toxoplasmosis scar, CNVM, macular detachment               |
| 20  | Diffuse unilateral subacute neuroretinitis, small worm     |
|   | Diffuse unilateral subacute neuroretinitis, large worm     |
|   | Cytomegalic inclusion disease, papillitis                  |
|   | Acute posterior multifocal placoid pigment epitheliopathy  |
|   | Acute macular neuroretinitis                               |
| 25  | Sarcoid retinitis  |
|   | Sarcoid papillitis   |
|   | Behcet's disease   |
|   | Vitiliginous (bird-shot) chorioretinitis                   |



-47-

|    |  |
|----|--|
|    | <b>Other Diseases</b>                                  |
|    | Multifocal choroiditis and panuveitis (pseudo-POHS)    |
|    | <b>Retinal and Pigment Epithelial Hamartomas</b>       |
|    | Congenital grouped albinotic RPE spots                 |
|    | Congenital hyperplasia of RPE                          |
| 5  | Combined RPE and retinal hamartoma, juxtapapillary     |
|    | Combined RPE and retinal hamartoma, peripheral         |
|    | Cystic astrocytoma, juxtapapillary                     |
|    | Astrocytoma, macula                                    |
|    | Astrocytoma, juxtapapillary                            |
| 10 | Cavernous hemangioma of retina                         |
|    | Juxtapapillary sessile retinal capillary hemangioma    |
|    | Juxtapapillary endophytic retinal capillary hemangioma |
|    | <b>Other Tumors of the Choroid</b>                     |
|    | Choroidal metastasis                                   |
| 15 | Choroidal osteoma                                      |
|    | Choroidal hemangioma                                   |
|    | Miscellaneous uveal tumors                             |
|    | <b>Intraocular Lymphoid Tumors</b>                     |
|    | The leukemias and lymphomas                            |
| 20 | <b>Tumors of the Vitreous</b>                          |
|    | Non-Hodgkins ("reticulum cell") lymphoma               |
|    | Tumor involvement of the vitreous cavity               |
|    | <b>Macular Disease</b>                                 |
|    | Age-related macular degeneration -- atrophic form      |
| 25 | Exudative age-related macular degeneration             |
|    | Choroidal neovascular membrane in degenerative myopia  |
|    | Central serous retinopathy                             |
|    | Macular hole   |

-48-

| Other Diseases           |   |
|--------------------------|---|
|                          | Macular dystrophies   |
| Retinal Vascular Disease |   |
|                          | Etiologic mechanisms in diabetic retinopathy                                |
|                          | Background diabetic retinopathy   |
| 5                        | Proliferative diabetic retinopathy  |
|                          | Retinal arterial obstructive disease  |
|                          | Central retinal vein occlusion  |
|                          | Retinal branch vein occlusion   |
|                          | Pregnancy and retinal disease   |
| 10                       | Pregnancy-induced hypertension  |
|                          | Hypertension  |
|                          | The rheumatic disease   |
|                          | Parafoveal telangiectasis   |
|                          | Coats disease   |
| 15                       | Disseminated intravascular systemic coagulopathy and related vasculopathies |
|                          | Hemoglobinopathies  |
|                          | Retinopathy of prematurity  |
|                          | Acquired retinal macroaneurysms   |
|                          | Eales disease   |
| 20                       | Radiation retinopathy   |
|                          | The ocular ischemic syndrome  |
| Inflammatory Disease     |   |
|                          | Ocular toxoplasmosis  |
|                          | Ocular toxocariasis   |
| 25                       | Ocular cysticercosis  |
|                          | Cytomegalovirus infections of the retina                                    |
|                          | Retinal and ophthalmologic manifestations of AIDS                           |
|                          | Acute retinal necrosis syndrome   |

-49-

| Other Diseases |   |
|----------------|---|
|                | Endogenous fungal infections of the retina and choroid                          |
|                | Pars planitis   |
|                | Syphilis and tuberculosis   |
|                | Diffuse unilateral subacute neuroretinitis                                      |
| 5              | Scleritis   |
|                | Birdshot retinochoroidopathy  |
|                | Punctate inner choroidopathy  |
|                | Sarcoidosis   |
|                | Acute multifocal placoid pigment epitheliopathy                                 |
| 10             | Geographic helicoid peripapillary choroidopathy (GHPC): serpiginous choroiditis |
|                | Sympathetic ophthalmia  |
|                | Vogt-Koyanagi-Harada syndrome (uveomeningitic syndrome)                         |
|                | Ciliochoroidal (uveal) effusion   |

15 Reproduced from: Stereoscopic Atlas of Ocular Diseases Diagnosis and Treatment, 2nd Edition, J. Donald O. Gass, Vol. 1 & 2, C.V. Mosley Co. (1987); and Retina Vol. II, Editor, Stephen J. Ryan, Medical Retina, C.V. Moslay Co. (1989).

## 20 2. Therapeutic products

Therapeutic products include but are not limited to, wild-type genes that are defective in ocular disorders, such as rhodopsin, or fragments thereof sufficient to correct the genetic defect, trophic factors, including growth factors, inhibitors and agonists of trophic factors, anti-apoptosis factors and other products described herein or known to those of skill in the art to be useful for treatment of disorders of the eye or that can be treated by a product expressed by a photoreceptor.

| OCULAR GENE THERAPY STRATEGIES |          |          |
|--------------------------------|----------|----------|
| GENERAL DISEASE                | EXAMPLES | STRATEGY |

-50-

| OCULAR GENE THERAPY STRATEGIES              |  |   |
|---|--|---|
| Hereditary retinal and macular degeneration | <ul style="list-style-type: none"> <li>• Retinitis pigmentosa</li> <li>• Stargardt's disease</li> <li>• Other macular dystrophies</li> </ul> | Growth factors (e.g., GDNF)<br>anti-apoptotic factors (e.g., bcl2 gene)<br>Stargardt Disease Gene (ABCR) <sup>†</sup> |
| Neovascular                                 | <ul style="list-style-type: none"> <li>• Diabetes</li> <li>• Choroidal neovascularization</li> </ul>   | Anti-angiogenesis factors   |
| Anti-tumor                                  | Retinoblastoma   | Antiproliferant   |
| 5    Glaucoma                               | Nerve fiber layer atrophy  | Neuroprotective agent   |

<sup>†</sup> See Allikmets *et al.* (1997) *Science* 277:1805-1807.

For example, for treatment of retinitis pigmentosa the adenovirus vector can deliver a wild-type rhodopsin gene or a growth factor or trophic factor, such as ciliary neurotrophic factor CNTF; for treatment of Stargardt's disease, the  
 10    vector can deliver a wild type ABCR (also called STGD1) or a growth factor or anti-angiogenic agent; for diabetic retinopathies, retinal vascularization the vector can deliver growth factors, such as a TGF (TGF $\beta$ ), to prevent degeneration.

15        The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

#### EXAMPLE 1

##### Preparation of Adenovirus Packaging Cell Lines

Cell lines that are commonly used for growing adenovirus are useful as  
 20    host cells for the preparation of adenovirus packaging cell lines. Preferred cells include 293 cells, an adenovirus-transformed human embryonic kidney cell line obtained from the ATCC, having Accession Number CRL 1573; HeLa, a human epithelial carcinoma cell line (ATCC Accession Number CCL-2); A549, a human lung carcinoma cell line (ATCC Accession Number CCL 1889); and other  
 25    epithelial-derived cell lines. As a result of the adenovirus transformation, the

-51-

293 cells contain the E1 early region regulatory gene. All cells were maintained in complete DMEM + 10% fetal calf serum unless otherwise noted.

These cell lines allow the production and propagation of adenovirus-based gene delivery vectors that have deletions in preselected gene regions and that are obtained by cellular complementation of adenoviral genes. To provide the desired complementation of such deleted adenoviral genomes in order to generate a viral vector, plasmid vectors that contain preselected functional units have been designed. Such units include but are not limited to E1 early region, E4 and the viral fiber gene. The preparation of plasmids providing such complementation, thereby being "complementary plasmids or constructs," that are stably inserted into host cell chromosomes are described below.

**A. Preparation of an E4-Expressing Plasmid for Complementation of E4-Gene-Deleted Adenoviruses**

The viral E4 regulatory region contains a single transcription unit that is alternately spliced to produce several different mRNA products. The E4-expressing plasmid prepared as described herein and used to transfect the 293 cell line contains the entire E4 transcription unit. A DNA fragment extending from 175 nucleotides upstream of the E4 transcription start site including the natural E4 promoter to 153 nucleotides downstream of the E4 polyadenylation signal including the natural E4 terminator signal, corresponding to nucleotides 32667-35780 of the adenovirus type 5 (hereinafter referred to as Ad5) genome as described in Chroboczek *et al.* (*Virology*, 186:280-285 (1992), GenBank Accession Number M73260), was amplified from Ad5 genomic DNA, obtained from the ATCC, via the polymerase chain reaction (PCR). Sequences of the primers used were 5'CGGTACACAGAATTCAGGAGACACAACCTCC3' (forward or 5' primer referred to as E4L) (SEQ ID NO: 1) and 5'GCCTGGATCCGGGAAGTTACGTAACGTGGGAAAAC3' (SEQ ID NO: 2) (backward or 3' primer referred to as E4R). To facilitate cloning of the PCR fragment, these oligonucleotides were designed to create new sites for the restriction enzymes EcoRI and BamHI, respectively, as indicated with underlined nucleotides. DNA was amplified via PCR using 30 cycles of 92 C for 1 minute,

-52-

50 C for 1 minute, and 72 C for 3 minutes resulting in amplified full-length E4 gene products.

The amplified DNA E4 products were then digested with EcoRI and BamHI for cloning into the compatible sites of pBluescript/SK+ by standard techniques to create the plasmid pBS/E4. A 2603 base pair (bp) cassette including the herpes simplex virus thymidine kinase promoter, the hygromycin resistance gene, and the thymidine kinase polyadenylation signal was excised from the plasmid pMEP4 (Invitrogen, San Diego, CA) by digestion with FspI followed by addition of BamHI linkers (5'CGCGGATCCGCG3') (SEQ ID NO: 3) for subsequent digestion with BamHI to isolate the hygromycin-containing fragment.

The isolated BamHI-modified fragment was then cloned into the BamHI site of pBS/E4 containing the E4 region to create the plasmid pE4/Hygro containing 8710 bp. The pE4/Hygro plasmid has been deposited with the ATCC under accession number 97739. The complete nucleotide sequence of pE4/Hygro is set forth in SEQ ID NO: 4. Position number 1 of the linearized vector corresponds to approximately the middle portion of the pBS/SK+ backbone. The 5' and 3' ends of the E4 gene are located at respective nucleotide positions 3820 and 707 of SEQ ID NO: 4 while the 5' and 3' ends of the hygromycin insert are located at respective nucleotide positions 3830 and 6470. In the clone that was selected for use, the E4 and hygromycin resistance genes were divergently transcribed.

**B. Preparation of a Fiber-Expressing Plasmid for Complementation of Fiber-Gene-Deleted Adenoviruses**

To prepare a fiber-encoding construct, primers were designed to amplify the fiber coding region from Ad5 genomic DNA with the addition of unique BamHI and NotI sites at the 5' and 3' ends of the fragment, respectively. The Ad5 nucleotide sequence is available with the GenBank Accession Number M18369. The 5' and 3' primers had the respective nucleotide sequences of 5'ATGGGATCCAAGATGAAGCGCGCAAGACCG3' (SEQ ID NO: 5) and 5'CATAACGCGGCCGCTTCTTTATTCTTGGGC3' (SEQ ID NO: 6), where the inserted BamHI and NotI sites are indicated by underlining. The 5' primer also

-53-

contained a nucleotide substitution 3 nucleotides 5' of the second ATG codon (C to A) that is the initiation site. The nucleotide substitution was included so as to improve the consensus for initiation of fiber protein translation.

The amplified DNA fragment was inserted into the BamHI and NotI sites of pcDNA3 (Invitrogen) to create the plasmid designated pCDNA3/Fiber having 7148 bp. The parent plasmid contained the CMV promoter, the bovine growth hormone (BHG) terminator and the gene for conferring neomycin resistance. The viral sequence included in this construct corresponds to nucleotides 31040-32791 of the Ad5 genome.

10 The complete nucleotide sequence of pCDNA3/Fiber is listed in SEQ ID NO: 7 where the nucleotide position 1 corresponds to approximately the middle of the pcDNA3 vector sequence. The 5' and 3' ends of the fiber gene are located at respective nucleotide positions 916 with ATG and 2661 with TAA.

To enhance expression of fiber protein by the constitutive CMV promoter provided by the pcDNA vector, a BglII fragment containing the tripartite leader (TPL) of adenovirus type 5 was excised from pRD112a (Sheay *et al.*, *BioTechniques*, 15:856-862 (1993) and inserted into the BamHI site of pCDNA3/Fiber to create the plasmid pCLF having 7469 bp. The adenovirus tripartite leader sequence, present at the 5' end of all major late adenoviral mRNAs as described by Logan *et al.*, *Proc. Natl. Acad. Sci., USA*, 81:3655-3659 (1984) and Berkner, *BioTechniques*, 6:616-629 (1988), also referred to as a "partial TPL" since it contains a partial exon 1, shows correspondence with the Ad5 leader sequence having three spatially separated exons corresponding to nucleotide positions 6081-6089 (the 3' end of the first leader segment), 7111-7182 (the entire second leader segment), and 9644-9845 (the third leader segment and sequence downstream of that segment). The corresponding cDNA sequence of the partial tripartite leader sequence present in pCLF is included in SEQ ID NO: 8 bordered by BamHI/BglII 5' and 3' sites at respective nucleotide positions 907-912 to 1228-1233. The nucleotide sequence of an isolated partial TPL is also listed separately as SEQ ID No. 22 with the noted 5' and 3' restriction sites and with the following nucleotide regions identified: 1-6 nt BglII site; 1-18 nt polylinker; 19-27 nt last 9 nt of the first leader segment (exon 1);

-54-

28-99 nt second leader segment (exon 2); 100-187 nt third leader segment (exon 3); 188-301 nt contains the nt sequence immediately following the third leader in the genome with an unknown function; and 322-327 nt BglII site.

The pCLF plasmid has been deposited with the ATCC as described in

- 5 Example 4. The complete nucleotide sequence of pCLF is listed in SEQ ID NO: 8 where the nucleotide position 1 corresponds to approximately the middle of the pcDNA3 parent vector sequence. The 5' and 3' ends of the Ad5 fiber gene are located at respective nucleotide positions 1237-1239 with ATG and 2980-2982 with TAA.

10 C. Generation of an Adenovirus Packaging Cell Line Carrying Plasmids Encoding Functional E4 and Fiber Proteins

- The 293 cell line was selected for preparing the first adenovirus packaging line as it already contains the E1 gene as prepared by Graham *et al.*, *J. Gen. Virol.*, 36:59-74 (1977) and as further characterized by Spector, *Virol.*, 15 130:533-538 (1983). Before electroporation, 293 cells were grown in RPMI medium + 10% fetal calf serum. Four x 10<sup>6</sup> cells were electroporated with 20 µg each of pE4/Hygro DNA and pCLF DNA using a BioRad GenePulser and settings of 300 V, 25 µF. DNA for electroporation was prepared using the Qiagen system according to the manufacturer's instructions (Bio-Rad, Richmond, 20 CA).

Following electroporation, cells were split into fresh complete DMEM + 10% fetal calf serum containing 200 µg/ml Hygromycin B (Sigma, St. Louis, MO).

- From expanded colonies, genomic DNA was isolated using the 25 "MICROTURBOGEN" system (Invitrogen) according to manufacturer's instructions. The presence of integrated E4 DNA was assessed by PCR using the primer pair E4R and ORF6L (5'TGCTTAAGCGGCCGCGAAGGAGA AGTCC3') (SEQ ID NO: 9), the latter of which is a 5' forward primer near adenovirus 5 open reading frame 6.

- 30 One clone, designated 211, was selected exhibiting altered growth properties relative to that seen in parent cell line 293. The 211 clone contained the product, indicating the presence of inserted DNA corresponding to most, if



-55-

not all, of the E4 fragment contained in the pE4/Hygro plasmid. The 211 cell line has been deposited with the ATCC as described in Example 4. This line was further evaluated by amplification using the primer pair E4L/E4R described above, and a product corresponding to the full-length E4 insert was detected.

- 5 Genomic Southern blotting was performed on DNA restricted with EcoRI and BamHI. The E4 fragment was then detected at approximately one copy/genome compared to standards with the EcoRI/BamHI E4 fragment as cloned into pBS/E4 for use as a labeled probe with the Genius system according to manufacturer's instructions (Boehringer Mannheim, Indianapolis, IN). In DNA from the 211 cell
- 10 line, the labeled internal fragment pE4/Hygro hybridized with the isolated E4 sequences. In addition, the probe hybridized to a larger fragment which may be the result of a second insertion event.

- Although the 211 cell line was not selected by neomycin resistance, thus indicating the absence of fiber gene, to confirm the lack of fiber gene, the 211
- 15 cell line was analyzed for expression of fiber protein by indirect immunofluorescence with an anti-fiber polyclonal antibody and a FITC-labeled anti-rabbit IgG (KPL) as secondary. No immunoreactivity was detected. Therefore, to generate 211 clones containing recombinant fiber genes, the 211 clone was expanded by growing in RPMI medium and subjected to additional
- 20 electroporation with the fiber-encoding pCLF plasmid as described above.

- Following electroporation, cells were plated in DMEM + 10% fetal calf serum and colonies were selected with 200  $\mu$ g/ml G418 (Gibco, Gaithersburg, MD). Positive cell lines remained hygromycin resistant. These candidate
- 25 sublines of 211 were then screened for fiber protein expression by indirect immunofluorescence as described above. The three sublines screened, 211A, 211B and 211R, along with a number of other sublines, all exhibited nuclear staining qualitatively comparable to the positive control of 293 cells infected with AdRSV $\beta$ gal (1 pfu/cell) and stained 24 hours post-infection.

- Lines positive for nuclear staining in this assay were then subjected to
- 30 Western blot analysis under denaturing conditions using the same antibody. Several lines in which the antibody detected a protein of the predicted molecular weight (62 kd for the Ad5 fiber protein) were selected for further study including

-56-

211A, 211B and 211R. The 211A cell line has been deposited with ATCC as described in Example 4.

Immunoprecipitation analysis using soluble nuclear extracts from these three cell lines and a semipreparative electrophoresis system demonstrated that the fiber protein expressed is in the functional trimeric form characteristic of the native fiber protein. The predicted molecular weight of a trimerized fiber is 186 kd. Under denaturing conditions, the trimeric form was destroyed resulting in detectable fiber monomers. Those clones containing endogenous E1, newly expressed recombinant E4 and fiber proteins were selected for use in complementing adenovirus gene delivery vectors having the corresponding adenoviral genes deleted as described in Example 2.

**D. Preparation of an E1-Expressing Plasmid for Complementation of E1-Gene-Deleted Adenoviruses**

In order to prepare adenoviral packaging cell lines other than those based on the E1-gene containing 293 cell line as described in Example 1C above, plasmid vectors containing E1 alone or in various combinations with E4 and fiber genes are constructed as described below.

The region of the adenovirus genome containing the E1a and E1b gene is amplified from viral genomic DNA by PCR as previously described. The primers used are E1L, the 5' or forward primer, and E1R, the 3' or backward primer, having the respective nucleotide sequences 5'CCG AGCTAGC GACTGAAAATGAG3' (SEQ ID NO: 10) and 5'CCTCTCGAG AGACAGC AAGACAC3' (SEQ ID NO: 11). The E1L and E1R primers include the respective restriction sites NheI and XhoI as indicated by the underlines. The sites are used to clone the amplified E1 gene fragment into the NheI/XhoI sites in pMAM commercially available from Clontech (Palo Alto, CA) to form the plasmid pDEX/E1 having 11152 bp.

The complete nucleotide sequence of pDEX/E1 is listed in SEQ ID NO: 12 where the nucleotide position 1 corresponds to approximately 1454 nucleotides from the 3' end of the pMAM backbone vector sequence. The pDEX/E1 plasmid includes nucleotides 552 to 4090 of the adenovirus genome positioned downstream (beginning at nucleotide position 1460 and ending at 4998 in the

-57-

pDEX/E1 plasmid) of the glucocorticoid-inducible mouse mammary tumor virus (MMTV) promoter of pMAM. The pMAM vector contains the E. coli *gpt* gene that allows stable transfectants to be isolated using hypoxanthine/amino-pterin/thymidine (HAT) selection. The pMAM backbone occupies nucleotide  
5 positions 1-1454 and 5005-11152 of SEQ ID NO: 12.

**E. Generation of an Adenovirus Packaging Cell Line Carrying Plasmids Encoding Functional E1, and Fiber Proteins**

To create separate adenovirus packaging cell lines equivalent to that of the 211 sublines, 211A, 211B and 211R, as described in Example 1C,  
10 alternative cell lines lacking adenoviral genomes are selected for transfection with the plasmid constructs as described below. Acceptable host cells include A549, Hela, Vero and the like cell lines as described in Example 1. The selected cell line is transfected with the separate plasmids, pDEX/E1 and pCLF, respectively for expressing E1, and fiber complementary proteins. Following  
15 transfection procedures as previously described, clones containing stable insertions of the two plasmids are isolated by selection with neomycin and HAT. Integration of full-length copy of the E1 gene is assessed by PCR amplification from genomic DNA using the primer set E1L/E1R, as described above. Functional insertion of the fiber gene is assayed by staining with the anti-fiber  
20 antibody as previously described.

The resultant stably integrated cell line is then used as a packaging cell system to complement adenoviral gene delivery vectors having the corresponding adenoviral gene deletions as described in Example 2.

**F. Preparation of a Plasmid Containing Two or More Adenoviral Genes for  
25 Complementing Gene-Deleted Adenoviruses**

The methods described in the preceding Examples rely on the use of two plasmids, pE4/Hygro and pCLF, or, pCLF and pDEX/E1 for generating adenoviral cell packaging systems. In alternative embodiments, complementing plasmids containing two or more adenoviral genes for expressing of encoded proteins in  
30 various combinations are also prepared as described below. The resultant plasmids are then used in various cell systems with delivery plasmids having the corresponding adenoviral gene deletions. The selection of packaging cell, content of the delivery plasmids and content of the complementing plasmids for

-58-

use in generating recombinant adenovirus viral vectors thus depends on whether other adenoviral genes are deleted along with the adenoviral fiber gene, and, if so, which ones.

5                   1.     **Preparation of a Complementing Plasmid Containing Fiber and E1 Adenoviral Genes**

A DNA fragment containing sequences for the CMV promoter, adenovirus tripartite leader, fiber gene and bovine growth hormone terminator is amplified from pCLF prepared in Example 1B using the forward primer 5'GACGGATCGGGAGATCTCC3' (SEQ ID NO: 13), that anneals to the  
10 nucleotides 1-19 of the pCDNA3 vector backbone in pCLF, and the backward primer 5'CCGCCTCAGAAGCCATAGAGCC3' (SEQ ID NO: 14) that anneals to nucleotides 1278-1257 of the pCDNA3 vector backbone. The fragment is amplified as previously described and then cloned into the pDEX/E1 plasmid, prepared in Example 1D. For cloning in the DNA fragment, the pDEX/E1 vector  
15 is first digested with NdeI, that cuts at a unique site in the pMAM vector backbone in pDEX/E1, then the ends are repaired by treatment with bacteriophage T4 polymerase and dNTPs.

The resulting plasmid containing E1 and fiber genes, designated pE1/Fiber, provides dexamethasone-inducible E1 function as described for  
20 DEX/E1 and expression of Ad5 fiber protein as described above.

The complete nucleotide sequence of pE1/Fiber is listed in SEQ ID NO: 15 where the nucleotide position 1 corresponds to approximately 1459 nucleotides from the 3' end of the parent vector pMAM sequence. The 5' and 3' ends of the Ad5 E1 gene are located at respective nucleotide positions 1460 and 4998  
25 followed by pMAM backbone and then separated from the Ad5 fiber from pCLF by the filled-in blunt ended NdeI site. The 5' and 3' ends of the pCLF fiber gene fragment are located at respective nucleotide positions 10922-14223 containing elements as previously described for pCLF.

The resultant pE1/Fiber plasmid is then used to complement one or more  
30 delivery plasmids expressing E1 and fiber.

The pE1/Fiber construct is then used to transfect a selected host cell as described in Example 1E to generate stable chromosomal insertions performed as

-59-

previously described followed by selection on HAT medium. The stable cells are then used as packaging cells as described in Example 2.

## **2. Preparation of a Complementing Plasmid Containing E4 and Fiber Adenoviral Genes**

5 Plasmid pCLF prepared as described in Example 1B is partially digested with BglII to cut only at the site in the pCDNA3 backbone. The pE4/Hygro plasmid prepared in Example 1A is digested with BamHI to produce a fragment containing E4. The E4 fragment is then inserted into the BamHI site of pCLF to form plasmid pE4/Fiber. The resultant plasmid provides expression of the fiber  
10 gene as described for pCLF and E4 function as described for pE4/Hygro.

A schematic plasmid map of pE4/Fiber, having 10610 bp. The complete nucleotide sequence of pE4/Fiber is listed in SEQ ID NO: 16 where the nucleotide position 1 corresponds to approximately 14 bp from the 3' end of the parent vector pCDNA3 backbone sequence. The 5' and 3' ends of the Ad5 E4  
15 gene are located at respective nucleotide positions 21 and 3149 followed by fused BglII/BamHI sites and pCDNA3 backbone including the CMV promoter again followed by BglII/BamHI sites. The adenovirus leader sequence begins at nucleotide position 4051 and extends to 4366 followed by fused BamHI/BglII sites and the 5' and 3' ends of the fiber gene located at respective nucleotide  
20 positions 4372 and 6124.

Stable chromosomal insertions of pE4/Fiber in host cells are obtained as described above.

## **EXAMPLE 2**

### **25 Preparation of Adenoviral Gene Delivery Vectors Using Adenoviral Packaging Cell Lines**

Adenoviral delivery vectors are prepared to separately lack the combinations of E1/fiber and E4/fiber. Such vectors are more replication-defective than those previously in use due to the absence of multiple viral genes. A preferred adenoviral delivery vector is replication competent but only via a  
30 non-fiber means is one that only lacks the fiber gene but contains the remaining functional adenoviral regulatory and structural genes. Furthermore, these adenovirus delivery vectors have a higher capacity for insertion of foreign DNA.

-60-

**A. Preparation of Adenoviral Gene Delivery Vectors Having Specific Gene Deletions and Methods of Use**

- To construct the E1/ fiber deleted viral vector containing the LacZ reporter gene construct, two new plasmids were constructed. The plasmid pΔ
- 5 E1Bβgal was constructed as follows. A DNA fragment containing the SV40 regulatory sequences and *E. coli* β-galactosidase gene was isolated from pSVβgal (Promega) by digesting with VspI, filling the overhanging ends by treatment with Klenow fragment of DNA polymerase I in the presence of dNTP's and digesting with Bam HI. The resulting fragment was cloned into the EcoRV and BamHI
- 10 sites in the polylinker of pΔ E1sp1B (Microbix Biosystems, Hamilton, Ontario) to form pΔ E1B βgal that therefore contained the left end of the adenovirus genome with the Ela region replaced by the LacZ cassette (nucleotides 6690 to 4151) of pSVβ gal. Plasmid DNA may be prepared by the alkaline lysis method as described by Birnboim and Doly, *Nuc. Acids Res.*, 7:1513-1523 (1978) or by the
- 15 Quiagen method according to the manufacturer's instruction, from transformed cells used to expand the plasmid DNA was then purified by CsCl-ethidium bromide density gradient centrifugation. Alternatively, plasmid DNAs may be purified from *E. coli* by standard methods known in the art (e.g. see Sambrook *et al.*)
- 20 The second plasmid (pDV44), prepared as described herein, is derived from pBHG10, a vector prepared as described by Bett *et al.*, *Proc. Natl. Acad. Sci., USA*, 91:8802-8806 (1994) (see, also International PCT application No. WO 95/00655) using methods well known to one of skill in the art. This vector is also commercially available from Microbix and contains an Ad5 genome
- 25 with the packaging signals at the left end deleted and the E3 region (nucleotides 28133:30818) replaced by a linker with a unique site for the restriction enzyme PacI. An 11.9 kb BamHI fragment, which contains the right end of the adenovirus genome, is isolated from pBHG10 and cloned into the BamHI site of pBS/SK(+) to create plasmid p11.3 having approximately 14,658 bp. The p11.3
- 30 plasmid was then digested with PacI and Sall to remove the fiber, E4, and inverted terminal repeat (ITR) sequences.

-61-

This fragment was replaced with a 3,4 kb fragment containing the ITR segments and the E4 gene which was generated by PCR amplification from pBHG10 using the following oligonucleotide sequences:

5' TGTACACCG GATCCGGCGCACACC3' SEQ ID NO: 17; and

5' CACAACGAGCTC AATTAATTAATTGCCACATCCTC3' SEQ ID NO: 18.

These primers incorporated sites for *PacI* and *BamHI*. Cloning this fragment into the *PacI* and blunt ended *Sall* sites of the p11.3 backbone resulted in a substitution of the fused ITRs, E4 region and fiber gene present in pBHG10, by the ITRs and E4 region alone. The resulting p11.3 plasmid containing the ITR and E4 regions, designated plasmid pDV43a, was then digested with *BamHI*. This *BamHI* fragment was then used to replace a *BamHI* fragment in pBHG10 thereby creating pDV44 in a pBHG10 backbone.

In an alternative approach to preparing pDV44 with an additional subcloning step to facilitate the incorporation of restriction cloning sites, the following cloning procedure was performed. pDV44 as above was constructed by removing the fiber gene and some of the residual E3 sequences from pBHG10 (Microbix Biosystems). As above, to simplify manipulations, the 11.9 kb *BamHI* fragment including the rightmost part of the Ad5 genome was removed from pBHG10 and inserted into pBS/SK. The resulting plasmid was termed p11.3. The 3.4 kb DNA fragment corresponding to the E4 region and both ITRs of adenovirus type 5 was amplified as described above from pBHG10 using the oligonucleotides listed above and subcloned into the vector pCR2.1 (Invitrogen) to create pDV42. This step is the additional cloning step to facilitate the incorporation of a *Sall* restriction site. pDV42 was then digested with *PacI*, which cuts at a unique site (bold type) in one of the PCR primers, and with *Sall*, which cuts at a unique site in the pCR2.1 polylinker. This fragment was used to replace the corresponding *PacI/XhoI* fragment of p11.3 (the pBS polylinker adjacent to the Ad DNA fragment contains a unique *XhoI* site), creating pDV43.

A plasmid designated pDV44 was constructed by replacing the 11.9 kb *BamHI* fragment of pBHG10 by the analogous *BamHI* fragment of pDV43. As generated in the first procedure, pDV44 therefore differs from pBHG10 by the

-62-

deletion of Ad5 nucleotides 30819:32743 (residual E3 sequences and all but the 3'-most 41 nucleotides of the fiber open reading frame).

Thus, to summarize, the cloning procedures described above result in the production of a fiber-deleted Ad5 genomic plasmid (pDV44) that was  
5 constructed by removing the fiber gene and some of the residual E3 sequences from pBHG10. pDV44 contains a wild-type E4 region, but only the last 41 nucleotides of the fiber ORF (this sequence was retained to avoid affecting expression of the adjacent E4 transcription unit). Plasmids pBHG10 and pDV44 contain unpackageable Ad5 genomes, and must be rescued by cotransfection  
10 and subsequent homologous recombination with DNA carrying functional packaging signals. In order to generate vectors marked with a reporter gene, either pDV44 or pBHG10 was cotransfected with p $\Delta$ E1B $\beta$ gal, which contains the left end of the Ad5 genome with an SV40-driven  $\beta$ -galactosidase reporter gene inserted in place of the E1 region.

15 In general, and as described below, the method for virus production by recombination of plasmids followed by complementation in cell culture involves the isolation of recombinant viruses by cotransfection of any one of the adenovirus packaging cell systems prepared in Example 1, namely 211A, 211B, 211R, A549, Vero cells, and the like, with plasmids carrying sequences  
20 corresponding to viral gene delivery vectors.

A selected cell line is plated in dishes and cotransfected with pDV44 and p $\Delta$ E1B $\beta$  gal using the calcium phosphate method as described by Bett *et al.*, *Proc. Natl. Acad. Sci., USA*, 91:8802-8806 (1994). Recombination between the overlapping adenovirus sequences in the two plasmids leads to the creation of a  
25 full-length viral chromosome where pDV44 and p $\Delta$ E1B $\beta$  gal recombine to form a recombinant adenovirus vector having multiple deletions. The deletion of E1 and of the fiber gene from the viral chromosome is compensated for by the sequences integrated into the packaging cell genome, and infectious virus particles are produced. The plaques thus generated are isolated and stocks of  
30 the recombinant virus are produced by standard methods.

Because of the fiber deletion, a pDV44-derived virus is replication-defective, cells in which it is grown must complement this defect.



-63-

The 211B cell line (a derivative of 293 cells which expresses the wild-type (wt) AD5 fiber and is equivalent to 211A on deposit with ATCC as described in Example 4) was used for rescue and propagation of the virus described here. pDV44 and pΔE1βgal were cotransfected into 211B cells, and the monolayers  
5 were observed for evidence of cytopathic effect (CPE). Briefly, for virus construction, cells were transfected with the indicated plasmids using the Gibco Calcium Phosphate Transfection system according to the manufacturer's instructions and observed daily for evidence of CPE.

One of a total of 58 transfected dishes showed evidence of spreading cell  
10 death at day 15. A crude freeze-thaw lysate was prepared from these cells and the resulting virus (termed Ad5.βgal.ΔF) was plaque purified twice and then expanded. To prepare purified viral preparations, cells were infected with the indicated Ad and observed for completion of CPE. Briefly, at day zero, 211B cells were plated in DMEM plus 10% fetal calf serum at approximately  $1 \times 10^7$   
15 cells/150 cm<sup>2</sup> flask or equivalent density. At day one, the medium was replaced with one half the original volume of fresh DMEM containing the indicated Ad, in this case Ad5.βgal.ΔF, at approximately 100 particles/cell. At day two, an equal volume of medium was added to each flask and the cells were observed for CPE. Two to five days after infection, cells were collected and virus isolated by lysis  
20 via four rapid freeze-thaw cycles. Virus was then purified by centrifugation on preformed 15-40% CsCl gradients ( $111,000 \times g$  for three hours at 4°C). The bands were harvested, dialyzed into storage buffer (10 mM Tris-pH 8.1, 0.9% NaCl, and 10% glycerol), aliquoted and stored at -70°C. Purified Ad5.βgal.ΔF virus particles containing human adenovirus Ad5.βgal.ΔF genome (described  
25 further below) have been deposited with the ATCC on January 15, 1999 as further described in Example 4.

For viral titering, as necessary in the below Examples, Ad preparations were titered by plaque assay on 211B cells. Cells were plated on  
polylysine-coated 6 well plates at  $1.5 \times 10^6$  cells/well. Duplicate dilutions of  
30 virus stock were added to the plates in 1 ml/well of complete DMEM. After a five hour incubation at 37°C, virus was removed and the wells overlaid with 2

-64-

ml of 0.6% low-melting agarose in Medium 199 (Gibco). An additional 1 ml of overlay was added at five day intervals.

As a control, the first-generation virus Ad5. $\beta$ gal.wt, which is identical to Ad5. $\beta$ gal. $\Delta$ F except for the fiber deletion, was constructed by cotransfection of pBHG10 and p $\Delta$ E1B $\beta$ gal. In contrast to the low efficiency of recovery of the fiberless genome (1/58 dishes), all of 9 dishes cotransfected with p $\Delta$ E1B $\beta$ gal and pBHG10 produced virus.

In another embodiment, a delivery plasmid is prepared that does not require the above-described recombination events to prepare a viral vector having a fiber gene deletion. In one embodiment, a single delivery plasmid containing all the adenoviral genome necessary for packaging but lacking the fiber gene is prepared from plasmid pFG140 containing full-length Ad5 that is commercially available from Microbix. The resultant delivery plasmid referred to as pFG140-f is then used with pCLF stably integrated cells as described above to prepare a viral vector lacking fiber. For genetic therapy, the fiber gene can be replaced with a therapeutic gene of interest for preparing a therapeutic delivery adenoviral vector. Methods for producing a fiberless vector with a complete TPL are described in Example 3.

Vectors for the delivery of any desired gene and preferably a therapeutic gene are prepared by cloning the gene of interest into the multiple cloning sites in the polylinker of commercially available p $\Delta$ E1sp1B (Microbix Biosystems), in an analogous manner as performed for preparing pE1B $\beta$ gal as described above. The same cotransfection and recombination procedure is then followed as described herein to obtain viral gene delivery vectors as further discussed in later Examples.

#### 1. Characterization of the Ad5. $\beta$ gal. $\Delta$ F Genome

To confirm that the vector genomes had the proper structures and that the fiber gene was absent from the Ad5. $\beta$ gal. $\Delta$ F chromosome, the DNA isolated from viral particles was analyzed. Briefly, purified viral DNA was obtained by adding 10  $\mu$ l of 10 mg/ml proteinase K, 40  $\mu$ l of 0.5 M EDTA and 50  $\mu$ l of 10% SDS to 800  $\mu$ l of adenovirus-containing culture supernatant. The suspension

-65-

was then incubated at 55°C for 60 minutes. The solution was then extracted once with 400  $\mu$ l of a 24:1 mixture of chloroform:isoamyl alcohol. The aqueous phase was then removed and precipitated with sodium acetate/ethanol. The pellet was washed once with 70% ethanol and lightly dried. The pellet was then suspended in 40  $\mu$ l of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA. Genomic DNA from Ad5. $\beta$ gal.wt and Ad5. $\beta$ gal. $\Delta$ F produced the expected restriction patterns following digestion with either EcoRI or with NdeI. Southern blotting, performed with standard methods, with labeled fiber DNA as a probe demonstrated the presence of fiber sequence in Ad5. $\beta$ gal.wt but not in Ad5. $\beta$ gal. $\Delta$ F DNA. As a positive control, the blot was stripped and reprobed with labeled E4 sequence. Fiber and E4 sequences were detected by using labeled inserts from pCLF and pE4/Hygro, respectively. E4 signal was readily detectable in both genomes at equal intensities. The complete nucleotide sequence of Ad5. $\beta$ gal. $\Delta$ F is presented in SEQ ID NO: 23 and is contained in the virus particle on deposit with ATCC.

## 2. Characterization of the Fiberless Adenovirus Ad5. $\beta$ gal. $\Delta$ F

To verify that Ad5. $\beta$ gal. $\Delta$ F was fiber-defective, 293 cells (which are permissive for growth of E1-deleted Ad vectors but do not express fiber) were infected with Ad5. $\beta$ gal. $\Delta$ F or with Ad5. $\beta$ gal.wt. Twenty-four hours post infection, the cells were stained with polyclonal antibodies directed either against fiber or against the penton base protein. Cells infected with either virus were stained by the anti-penton base antibody, while only cells infected with the Ad5. $\beta$ gal.wt control virus reacted with the anti-fiber antibody. This confirms that the fiber-deleted Ad mutant does not direct the synthesis of fiber protein.

## 3. Growth of the Fiber-Deleted Ad5. $\beta$ gal. $\Delta$ F Vector in Complementing Cells

Ad5. $\beta$ gal. $\Delta$ F was found to readily be propagated in 211B cells. As assayed by protein concentration, CsCl-purified stocks of either Ad5. $\beta$ gal. $\Delta$ F or Ad5. $\beta$ gal.wt contained similar numbers of viral particles. The particles appeared to band normally on CsCl gradients. Infectivity of the Ad5. $\beta$ gal. $\Delta$ F particles was lower than the Ad5. $\beta$ gal.wt control, as indicated by an increased particle/PFU ratio. Ad5. $\beta$ gal. $\Delta$ F was also found to plaque more slowly than the control

-66-

virus. When plated on 211B cells, Ad5. $\beta$ gal.wt plaques appeared within 5-7 days, while plaques of Ad5. $\beta$ gal. $\Delta$ F continued to appear until as much as 15-18 days post infection. Despite their slower formation, the morphology of Ad5. $\beta$ gal. $\Delta$ F plaques was essentially normal.

5                                    4.        Production of Fiberless Ad5. $\beta$ gal. $\Delta$ F Particles

As Ad5. $\beta$ gal. $\Delta$ F represents a true fiber null mutation and its stocks are free of helper virus, the fiber mutant phenotype was readily investigated. A single round of growth in cells (such as 293) which do not produce fiber generating a homogeneous preparation of fiberless Ad allowed for the  
10        determination of whether such particles would be stable and/or infectious. Either Ad5. $\beta$ gal.wt or Ad5. $\beta$ gal. $\Delta$ F was grown in 293 or 211B cells, and the resulting particles purified on CsCl gradients as previously described. Ad5. $\beta$ gal. $\Delta$ F particles were readily produced in 293 cells at approximately the same level as the control virus and behaved similarly on the gradients, indicating  
15        that there was not a gross defect in morphogenesis of fiberless capsids.

Particles of either virus contained similar amounts of penton base regardless of the cell type in which they were grown. This demonstrated that fiber is not required for assembly of the penton base complex into virions. The Ad5. $\beta$ gal. $\Delta$ F particles produced in 293 cells did not contain fiber protein.  
20        211B-grown Ad5. $\beta$ gal. $\Delta$ F also contained less fiber than the Ad5. $\beta$ gal.wt control virus. The infectivities of the different viral preparations on epithelial cells correlated with the amount of fiber protein present. The fiberless Ad particles were several thousand-fold less infectious than the first-generation vector control on a per-particle basis, while infectivity of 211B-grown Ad5. $\beta$ gal. $\Delta$ F was only  
25        50-100 fold less than that of Ad5. $\beta$ gal.wt. These studies confirmed fiber's crucial role in infection of epithelial cells via CAR binding.

5.        Composition and Structure of the Fiberless Ad5. $\beta$ gal. $\Delta$ F Particles

The proteins contained in particles of 293-grown Ad5. $\beta$ gal. $\Delta$ F were  
30        compared to those in Ad5. $\beta$ gal.wt, to determine whether proteolysis or particle assembly was defective in this fiber null mutant. The overall pattern of proteins in the fiberless particles was observed to be quite similar to that of a

-67-

first-generation vector, with the exception of reduced intensity of the composite band resulting from proteins IIIa and IV (fiber). The fiberless particles also had a reduced level of protein VII. Although substantial amounts of uncleaved precursors to proteins VI, VII, and VIII were not seen, it is possible that the

5 low-molecular weight bands migrating ahead of protein VII represent either aberrantly cleaved viral proteins or their breakdown products.

Cryo-electron microscopy was used to more closely examine the structure of the 293 grown Ad5.βgal.ΔF and of Ad5βgal.wt. The fiber, having an extended stalk with a knob at the end, was faintly visible in favorable

10 orientations of wild-type Ad5 particles, but not in images of the fiberless particles. Filamentous material likely corresponding to free viral DNA was seen in micrographs of fiberless particles. This material was also present in micrographs of the first-generation control virus, albeit at much lower levels.

Three-dimensional image reconstructions of fiberless and wild-type

15 particles at ~20 Å resolution showed similar sizes and overall features, with the exception that fiberless particles lacked density corresponding to the fiber protein. The densities corresponding to other capsid proteins, including penton base and proteins IIIa, VI, and IX, were comparable in the two structures. This confirms that absence of fiber does not prevent assembly of these components

20 into virions. The fiber was truncated in the wild-type structure as only the lower portion of its flexible shaft follows icosahedral symmetry. The RGD protrusions on the fiberless penton base were angled slightly inward relative to those of the wild-type structure. Another difference between the two penton base proteins was that there is a ~30 Å diameter depression in the fiberless penton base

25 around the five-fold axis where the fiber would normally sit. The Ad5 reconstructions confirm that capsid assembly, including addition of penton base to the vertices, is able to proceed in the complete absence of fiber.

#### 6. Integrin-Dependent Infectivity of Fiberless Ad5.βgal.ΔF Particles

30 While attachment via the viral fiber protein is a critical step in the infection of epithelial cells, an alternative pathway for infection of certain hematopoietic cells has been described. In this case, penton base mediates

-68-

binding to the cells (via  $\beta 2$  integrins) and internalization (through interaction with  $\alpha v$  integrins). Particles lacking fiber might therefore be expected to be competent for infection of these cells, even though on a per-particle basis they are several thousand-fold less infectious than normal Ad vectors on epithelial cells.

To investigate this, THP-1 monocytic cells were infected with Ad5. $\beta$ gal:wt or with Ad5. $\beta$ gal. $\Delta$ F grown in the absence of fiber. Infection of THP-1 cells was assayed by infecting  $2 \times 10^5$  cells at the indicated m.o.i. in 0.5 ml of complete RPMI. Forty-eight hours post-infection, the cells were fixed with glutaraldehyde and stained with X-gal, and the percentage of stained cells was determined by light microscopy. The results of the infection assay showed that the fiberless particles were only a few-fold less infectious than first-generation Ad on THP-1 cells. Large differences were seen in plaquing efficiency on epithelial (211B) cells. Infection of THP-1 cells by either Ad5. $\beta$ gal. $\Delta$ F or Ad5. $\beta$ gal:wt was not blocked by an excess of soluble recombinant fiber protein, but could be inhibited by the addition of recombinant penton base). These results indicate that the fiberless Ad particles use a fiber-independent pathway to infect these cells. Furthermore, the lack of fiber protein did not prevent Ad5. $\beta$ gal. $\Delta$ F from internalizing into the cells and delivering its genome to the nucleus, demonstrating that fiberless particles are properly assembled and are capable of uncoating.

The foregoing results with the recombinant viruses thus produced indicates that they can be used as gene delivery tools in cultured cells and *in vivo* as described more fully in the Examples. For example, for studies of the effectiveness and relative immunogenicity of multiply-deleted vectors, virus particles are produced by growth in the packaging lines described in Example 1 and are purified by CsCl gradient centrifugation. Following titering, virus particles are administered to mice via systemic or local injection or by aerosol delivery to lung. The LacZ reporter gene allows the number and type of cells which are successfully transduced to be evaluated. The duration of transgene expression is evaluated in order to determine the long-term effectiveness of treatment with multiply-deleted recombinant adenoviruses relative to the

-69-

standard technologies which have been used in clinical trials to date. The immune response to the improved vectors described here is determined by assessing parameters such as inflammation, production of cytotoxic T lymphocytes directed against the vector, and the nature and magnitude of the antibody response directed against viral proteins.

Versions of the vectors which contain therapeutic genes such as CFTR for treatment of cystic fibrosis or tumor suppressor genes for cancer treatment are evaluated in the animal system for safety and efficiency of gene transfer and expression. Following this evaluation, they are used as experimental therapeutic agents in human clinical trials.

**B. Retargeting of Adenoviral Gene Delivery Vectors by Producing Viral Particles Containing Different or Altered Fiber Proteins**

As the specificity of adenovirus binding to target cells is largely determined by the fiber protein, viral particles that incorporate modified fiber proteins or fiber proteins from different adenoviral serotypes (pseudotyped vectors) have different specificities. Thus, the methods of expression of the native Ad5 fiber protein in adenovirus packaging cells as described above is also applicable to production of different fiber proteins.

Chimeric fiber proteins can be produced according to known methods (see, *e.g.*, Stevenson *et al.* (1995) *J. Virol.*, 69:2850-2857). Determinants for fiber receptor binding activity are located in the head domain of the fiber and an isolated head domain is capable of trimerization and binding to cellular receptors. The head domains of adenovirus type 3 (Ad3) and Ad5 were exchanged in order to produce chimeric fiber proteins. Similar constructs for encoding chimeric fiber proteins for use in the methods herein are contemplated. Thus, instead of the using the intact Ad5 fiber-encoding construct prepared in above and in U.S. application Serial No. 09/482,682) as a complementing viral vector in adenoviral packaging cells, the constructs described herein are used to transfect cells along with E4 and/or E1-encoding constructs.

Briefly, full-length Ad5 and Ad3 fiber genes were amplified from purified adenovirus genomic DNA as a template. The Ad5 and Ad3 nucleotide sequences are available with the respective GenBank Accession Numbers

-70-

M18369 and M12411. Oligonucleotide primers are designed to amplify the entire coding sequence of the full-length fiber genes, starting from the start codon, ATG, and ending with the termination codon TAA. For cloning purposes, the 5' and 3' primers contain the respective restriction sites BamHI and NotI for  
5 cloning into pcDNA plasmid as described in Example 1A. PCR is performed as described above.

The resulting products are then used to construct chimeric fiber constructs by PCR gene overlap extension (Horton *et al.* (1990) *BioTechniques*, 8:525-535). The Ad5 fiber tail and shaft regions (5TS; the nucleotide region  
10 encoding amino acid residue positions 1 to 403) are connected to the Ad3 fiber head region (3H; the nucleotide region encoding amino acid residue positions 136 to 319) to form the 5TS3H fiber chimera. Conversely, the Ad3 fiber tail and shaft regions (3TS; the nucleotide region encoding amino acid residues positions 1 to 135) are connected to the Ad5 fiber head region (5H; the nucleotide region  
15 encoding the amino acid residue positions 404 to 581) to form the 3TS5H fiber chimera. The fusions are made at the conserved TLWT (SEQ ID NO: 19) sequence at the fiber shaft-head junction.

The resultant chimeric fiber PCR products are then digested with BamHI and NotI for separate directional ligation into a similarly digested pcDNA 3.1.  
20 The TPL sequence is then subcloned into the BamHI as described in Example 1A for preparing an expression vector for subsequent transfection into 211 cells as described above or into the alternative packaging cell systems as previously described. The resultant chimeric fiber construct-containing adenoviral packaging cell lines are then used to complement adenoviral delivery vectors as  
25 previously described. Other fiber chimeric constructs are obtained with the various adenovirus serotypes using a similar approach.

In an alternative embodiment, the use of modified proteins including with modified epitopes (see, *e.g.*, Michael *et al.* (1995) *Gene Therapy*, 2:660-668 and International PCT application Publication No. WO 95/26412, which describe  
30 the construction of a cell-type specific therapeutic viral vector having a new binding specificity incorporated into the virus concurrent with the destruction of the endogenous viral binding specificity). In particular, the authors described the



-71-

production of an adenoviral vector encoding a gastrin releasing peptide (GRP) at the 3' end of the coding sequence of the Ad5 fiber gene. The resulting fiber-GRP fusion protein was expressed and shown to assemble functional fiber trimers that were correctly transported to the nucleus of HeLa cells following synthesis.

Similar constructs are contemplated for use in the complementing adenoviral packaging cell systems for generating new adenoviral gene delivery vectors that are targetable, replication-deficient and less immunogenic. Heterologous ligands contemplated for use herein to redirect fiber specificity range from as few as 10 amino acids in size to large globular structures, some of which necessitate the addition of a spacer region so as to reduce or preclude steric hindrance of the heterologous ligand with the fiber or prevent trimerization of the fiber protein. The ligands are inserted at the end or within the linker region. Preferred ligands include those that target specific cell receptors or those that are used for coupling to other moieties such as biotin and avidin.

A preferred spacer includes a short 12 amino acid peptide linker composed of a series of serines and alanine flanked by a proline residue at each end using routine procedures known to those of skill in the art. The skilled artisan will be with the preparation of linkers to accomplish sufficient protein presentation and to alter the binding specificity of the fiber protein without compromising the cellular events that follow viral internalization. Moreover, within the context of this disclosure, preparation of modified fibers having ligands positioned internally within the fiber protein and at the carboxy terminus as described below are contemplated for use with the methods described herein.

The preparation of a fiber having a heterologous binding ligand is prepared essentially as described in the above-cited paper. Briefly, for the ligand of choice, site-directed mutagenesis is used to insert the coding sequence for a linker into the 3' end of the Ad5 fiber construct in pCLF as prepared in Example

1.

The 3' or antisense or mutagenic oligonucleotide encodes a preferred linker sequence of ProSerAlaSerAlaSerAlaSerAlaProGlySer (SEQ ID NO: 20)

-72-

followed by a unique restriction site and two stop codons, respectively, to allow the insertion of a coding sequence for a selected heterologous ligand and to ensure proper translation termination. Flanking this linker sequence, the mutagenic oligonucleotide contains sequences that overlap with the vector sequence and allow its incorporation into the construct. Following mutagenesis of the pCLF sequence adding the linker and stop codon sequences, a nucleotide sequence encoding a preselected ligand is obtained, linkers corresponding to the unique restriction site in the modified construct are attached and then the sequence is cloned into linearized corresponding restriction site. The resultant fiber-ligand construct is then used to transfect 211 or the alternative cell packaging systems previously described to produce complementing viral vector packaging systems.

In a further embodiment, intact fiber genes from different Ad serotypes are expressed by 211 cells or an alternative packaging system as previously described. A gene encoding the fiber protein of interest is first cloned to create a plasmid analogous to pCLF, and stable cell lines producing the fiber protein are generated as described above for Ad5 fiber. The adenovirus vector described which lacks the fiber gene is then propagated in the cell line producing the fiber protein relevant for the purpose at hand. As the only fiber gene present is the one in the packaging cells, the adenoviruses produced contain only the fiber protein of interest and therefore have the binding specificity conferred by the complementing protein. Such viral particles are used in studies such as those described above to determine their properties in experimental animal systems.

### EXAMPLE 3

Tripartite leader sequences (TPLs) that are useful in enhancing the expression of complementing adenoviral proteins, particularly fiber protein, for use in preparing an adenoviral gene delivery vector are provided. The complete Ad5 TPL was constructed by assembling PCR fragments. First, the third TPL exon (exon 3) (nt 9644-9731 of the Ad5 genome) was amplified from Ad5 genomic DNA using the synthetic oligonucleotide primers 5'CTCAACAATTGTGGATCCGTACTCC3' (SEQ ID No. 24) and 5'GTGCTCAGCAGATCTTGCGACTGTG3' (SEQ ID No. 25). The resulting

-73-

product was cloned to the BamHI and BglII sites of pΔE1Sp1a (Microbix Biosystems) using sites in the primers (shown in bold) to create plasmid pDV52. A fragment corresponding to the first TPL exon (exon 1), the natural first intron (intron 1), and the second TPL exon (exon 2) (Ad5 nt 6049-7182) was then  
5 amplified using primers 5'GGCGCGTTCGGATCCACTCTCTTCC3' (SEQ ID No. 26) and 5'CTACATGCTAGGCAGATCTCGTTCGGAG3' (SEQ ID No. 27), and cloned into the BamHI site of pDV52 (again using sites in the primers) to create pDV55.

This plasmid contains a 1.2 kb BamHI/BglII fragment containing the first  
10 TPL exon, the natural first intron, and the fused second and third TPL exons. The nucleotide sequence of the complete TPL containing the noted 5' and 3' restriction sites is shown in SEQ ID No 28 with the following nucleotide regions identified: 1-6 nt BamHI site; 7-47 nt first leader segment (exon 1); 48-1068 nt natural first intron (intron 1); 1069-1140 nt second leader segment (exon 2);  
15 1141-1146 nt fused BamHI and BglII sites; 1147-1234 nt third leader segment (exon 3); and 1235-1240 nt BglII site.

#### EXAMPLE 4

##### Deposit of Materials

The following cell lines and plasmids were deposited on September 25,  
20 1996, with the American Type Culture Collection, 10801 University Blvd, Manassas, Virginia, USA (ATCC) under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty): Plasmid pE4/Hygro (accession number 97739), Plasmid pCLF (accession number 97737),  
25 211 Cell Line (accession number CRL-12193) and 211A Cell Line (accession number CRL-12194)

The following virus, Ad5.βgal.ΔF, was deposited on January 15, 1999, with the ATCC as listed above and provided with accession number VR2636.

Additionally, plasmids pDV60, pDV67, pDV69, pDV80 and pDV90 were  
30 deposited at the ATCC on January 5, 2000 and provided with accession numbers PTA-1144, PTA-1145, PTA-1146, PTA-1147 and PTA-1148 respectively.

-74-

**EXAMPLE 5****Preparation and Use of Adenoviral Packaging Cell Lines  
Containing Plasmids Containing Alternative TPLs**

5 Plasmids containing tripartite leaders (TPLs) have been constructed. The resulting plasmids that contain different selectable markers, such as neomycin and zeocin, were then used to prepare fiber-complementing stable cell lines for use as for preparing adenoviral vectors.

**A. pDV60**

10 Plasmid pDV60 was constructed by inserting this TPL cassette of SEQ ID No. 28 into the BamHI site upstream of the Ad5 fiber gene in pcDNA3/Fiber, a neomycin selectable plasmid (see, *e.g.*, U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000); see also Von Seggern *et al.* (1998) *J. Gen Virol.*, 79: 1461-1468). The nucleotide sequence of pDV60 is  
15 listed in SEQ ID NO: 29. Plasmid pDV60 has been deposited in the ATCC under accession number PTA-1144.

**B. pDV61**

To construct pDV61, an Asp718/NotI fragment containing the CMV promoter, partial Ad5 TPL, wildtype Ad5 fiber gene, and bovine growth hormone  
20 terminator was transferred from pCLF (ATCC accession number 97737; and described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000);), to a zeocin selectable cloning vector referred to as pCDNA3.1/Zeo (+) (commercially available from Invitrogen and for which the sequence is known).

**25 C. pDV67**

In an analogous process, pDV67 containing complete TPL was constructed by transferring an Asp 718/XbaI fragment from pDV60 into pcDNA3.1/Zeo(+) backbone. The nucleotide sequence of pDV67 is set forth in

-75-

SEQ ID No. 30. Plasmid pDV67 is available from the ATCC under accession number PTA-1145.

**D. pDV69**

To prepare pDV69 containing a modified fiber protein, the chimeric  
5 Ad3/Ad5 fiber gene was amplified from pGEM5TS3H (Stevenson *et al.* (1995) *J. Virol.*, 69: 2850-2857) using the primers 5'ATGGGAT  
CAAGATGAAGCGCGCAAGACCG3' (SEQ ID No. 31) and  
5'CACTATAGCGGCCGCATTCTCAGTCATCTT3' (SEQ ID No. 32), and cloned to  
the BamHI and NotI sites of pcDNA3.1/Zeo(+) via new BamHI and NotI sites  
10 engineered into the primers to create pDV68. Finally, the complete TPL  
fragment described above was then added to the unique BamHI site of pDV68  
to create pDV69. The nucleotide sequence of pDV69 is listed in SEQ ID No. 33  
and has been deposited in the ATCC under accession number PTA-1146.

**E. Preparation of Stable Adenovirus Packaging Cell Lines**

15 E1-2a S8 cells are derivatives of the A549 lung carcinoma line (ATCC #  
CCL 185) with chromosomal insertions of the plasmids pGRE5-2.E1 (also  
referred to as GRE5-E1-SV40-Hygro construct and listed in SEQ ID No. 34) and  
pMNeoE2a-3.1 (also referred to as MMTV-E2a-SV40-Neo construct and listed in  
SEQ ID No. 35), which provide complementation of the adenoviral E1 and E2a  
20 functions, respectively. This line and its derivatives were grown in Richter's  
modified medium (BioWhitaker) + 10% FCS. E1-2a S8 cells were  
electroporated as previously described (Von Seggern *et al.* (1998) *J. Gen Virol.*,  
79: 1461-1468) with pDV61, pDV67, or with pDV69, and stable lines were  
selected with zeocin (600 µg/ml).

25 The cell line generated with pDV61 is designated 601. The cell line  
generated with pDV67 is designated 633 while that generated with pDV69 is  
designated 644. Candidate clones were evaluated by immunofluorescent  
staining with a polyclonal antibody raised against the Ad2 fiber. Lines  
expressing the highest level of fiber protein were further characterized.

30 For the S8 cell complementing cell lines, to induce E1 expression, 0.3 µM  
of dexamethasone was added to cell cultures 16-24 hours prior to challenge  
with virus for optimal growth kinetics. For preparing viral plaques, 5 X 10<sup>5</sup>

-76-

cells/well in 6 well plates are prepared and pre-induced with the same concentration of dexamethasone the day prior to infection with 0.5  $\mu$ M included at a final concentration in the agar overlay after infection.

**F. Development of Cell Lines for Complementation of E1'/E2a' Vectors**

5       The Adenovirus 5 genome was digested with Scal enzyme, separated on an agarose gel, and the 6,095 bp fragment containing the left end of the virus genome was isolated. The complete Adenovirus 5 genome is registered as Genbank accession #M73260, incorporated herein by reference, and the virus is available from the American Type Culture Collection, Manassas, Virginia, U.S.A.,  
10       under accession number VR-5. The Scal 6,095 bp fragment was digested further with Clal at bp 917 and BglII at bp 3,328. The resulting 2,411 bp Clal to BglII fragment was purified from an agarose gel and ligated into the superlinker shuttle plasmid pSE280 (Invitrogen, San Diego, CA), which was digested with Clal and BglII, to form pSE280-E.

15       Polymerase chain reaction (PCR) was performed to synthesize DNA encoding an XhoI and SalI restriction site contiguous with Adenovirus 5 DNA bp 552 through 924. The primers which were employed were as follows:

5' end, Ad5 bp 552-585:

5'-GTCACTCGAGGACTCGGTC-GACTGAAAATGAGACATATTATCTGCCACGGA

20       CC-3' (SEQ ID No 36)

3' end, Ad5 bp 922-891:

5'-CGAGATCGATCACCTCCGGTACAAGGTTTGGCATAG-3' (SEQ ID No. 37)

      This amplified DNA fragment (sometimes hereinafter referred to as Fragment A) then was digested with XhoI and Clal, which cleaves at the native  
25       Clal site (bp 917), and ligated to the XhoI and Clal sites of pSE280-E, thus reconstituting the 5' end of the E1 region beginning 8 bp upstream of the ATG codon.

      PCR then was performed to amplify Adenovirus 5 DNA from bp 3,323 through 4,090 contiguous with an EcoRI restriction site. The primers which  
30       were employed were as follows:

5' end, Ad5 bp 3323-3360:

-77-

5'-CATGAAGATCTGGAAGGTGCTGAGGTACGATGAGACC-3' (SEQ ID No. 38);

and

3' end, Ad5 bp 4090-4060:

5'-GCGACTTAAGCAGTCAGCTG-AGACAGCAAGACACTTGCTTGATCCAAATCC

5 -3' (SEQ ID No. 39).

This amplified DNA fragment (sometimes hereinafter referred to as Fragment B) was digested with BglII, thereby cutting at the Adenovirus 5 BglII site (bp 3,382) and EcoRI, and ligated to the BglII and EcoRI sites of pSE280-AE to reconstruct the complete E1a and E1b region from Adenovirus 5 bp 552 through 4,090. The resulting plasmid is designated pSE280-E1.

A construct containing the intact E1a/b region under the control of the synthetic promoter GRE5 was prepared as follows. The intact E1a/b region was excised from pSE280-E1, which was modified previously to contain a BamHI site 3' to the E1 gene, by digesting with XhoI and BamHI. The XhoI to BamHI fragment containing the E1a/b fragment was cloned into the unique XhoI and BamHI sites of pGRE5-2/EBV (U.S. Biochemicals, Cleveland, Ohio) to form pGRE5-E1).

Bacterial transformants containing the final construct were identified. Plasmid DNA was prepared and purified by banding in CsTFA prior to use for transfection of cells.

**Construction of plasmid including Adenovirus 5 E2A sequence.**

The Adenovirus 5 genome was digested with BamHI and SpeI, which cut at bp 21,562 and 27,080, respectively. Fragments were separated on an agarose gel and the 5,518 bp BamHI to SpeI fragment was isolated. The 5,518 bp BamHI to SpeI fragment was digested further with SmaI, which cuts at bp 23,912. The resulting 2,350 bp BamHI to SmaI fragment was purified from an agarose gel, and ligated into the superlinker shuttle plasmid pSE280, and digested with BamHI and SmaI to form pSE280-E2 BamHI-SmaI.

PCR then was performed to amplify Adenovirus 5 DNA from the SmaI site at bp 23,912 through 24,730 contiguous with NheI and EcoRI restriction sites.

The primers which were employed were as follows:

5' end, Ad5 bp 24,732-24,708:

-78-

5'-CACGAATTCGTCAGCGCTTCTCGTCGCGTCCAAGACCC-3' (SEQ ID No. 40);  
3' end, Ad5 bp 23,912-23,934:

5'-CACCCCGGGGAGGCGGCGGCGACGGGGACGGG-3' (SEQ ID No. 41)

This amplified DNA fragment was digested with SmaI and EcoRI, and  
5 ligated to the SmaI and EcoRI sites of pSE280-E2 Bam-Sma to reconstruct the complete E2a region from Ad5 bp 24,730 through 21,562. The resulting construct is pSE280-E2a.

In order to convert the BamHI site at the 3' end of E2a to a Sall site, the E2a region was excised from pSE280-E2a by cutting with BamHI and NheI, and  
10 recloned into the unique BamHI and NheI sites of pSE280. Subsequently, the E2a region was excised from this construction with NheI and Sall in order to clone into the NheI and Sall sites of the pMAMneo (Clonetech, Palo Alto, CA) multiple cloning site in a 5' to 3' orientation, respectively. The resulting construct is pMAMneo E2a.

15 Bacterial transformants containing the final pMAMneo-E2a were identified. Plasmid DNA was prepared and purified by banding in CsTFA. Circular plasmid DNA was linearized at the XmnI site within the ampicillin resistance gene of pMAMneo-E2a, and further purified by the phenol/chloroform extraction and ethanol precipitation prior to use for transfection of cells.

#### 20 Transfection and selection of cells.

In general, this process involved the sequential introduction, by calcium phosphate precipitation, or other means of DNA delivery, of two plasmid constructions each with a different viral gene, into a single tissue culture cell. The cells were transfected with a first construct and selected for expression of  
25 the associated drug resistance gene to establish stable integrants. Individual cell clones were established and assayed for function of the introduced viral gene. Appropriate candidate clones then were transfected with a second construct including a second viral gene and a second selectable marker. Transfected cells then were selected to establish stable integrants of the second construct, and  
30 cell clones were established. Cell clones were assayed for functional expression of both viral genes.



-79-

A549 (ATCC Accession No. CCL-185) were used for transfection.

Appropriate selection conditions were established for G418 and hygromycin B by standard kill curve determination.

Transfection of A549 cells with plasmids including E1 and E2a regions.

- 5 pMAMNeo-E2a was linearized with XmnI with the Amp<sup>R</sup> gene, introduced into cells by transfection, and cells were selected for stable integration of this plasmid by G418 selection until drug resistant colonies arose. The clones were isolated and screened for E2a expression by staining for E2a protein with a polyclonal antiserum, and visualizing by immunofluorescence. E2a function was  
10 screened by complementation of the temperature-sensitive mutant Ad5ts125 virus which contains a temperature-sensitive mutation in the E2a gene. (Van Der Vliet, et al., J. Virology, Vol. 15, pgs. 348-354 (1975)). Positive clones expressing the E2a gene were identified and used for transfection with the 7 kb EcoRV to XmnI fragment from pGRE5-E1, which contains the GRE5 promoted  
15 E1a/b region plus the hygromycin<sup>R</sup> gene. Cells were selected for hygromycin resistance and assayed for E1a/b expression by staining with a monoclonal antibody for the E1 protein (Oncogene Sciences, Uniondale, N.Y.). E1 function was assayed by ability to complement an E1-deleted vector. At this point, expression and function of E2a was verified as described above, thus  
20 establishing the expression of E1a/b and E2a in the positive cell clones.

A transfected A549 ( A549 (ATCC Accession No. CCL-185);) cell lines showed good E1a/b and E2a expression and was selected for further characterization. It was designated the S8 cell line.

- 25 G. Preparation of Adenoviral Vectors Containing Ad5.βgal.ΔF Genome in S8 Improved Fiber-Complementing Cell Lines

- To prepare adenoviral vectors containing Ad5.βgal.ΔF (Ad5.βgal.ΔF has been deposited the ATCC under accession number VR2636) in S8 cells containing alternative forms of TPL for enhancing the expression of fiber proteins, the protocol as described in Example 2 for preparing Ad5.βgal.ΔF in  
30 211B cells was followed with the exception of pretreatment with 0.3 μM dexamethasone for 24 hours as described above. Thus, viral particles with the wildtype Ad5 fiber protein on their surface and containing the fiberless

-80-

Ad5. $\beta$ gal. $\Delta$ F genome were produced in 633 cells. Particles produced in 644 cells also contained the fiberless Ad5. $\beta$ gal. $\Delta$ F genome, but had the chimeric 5T3H fiber protein, with the Ad3 fiber knob, on their surface.

Thus, these viral preparations, prepared as described herein are useful for targeting delivery of the Ad5. $\beta$ gal. $\Delta$ F, Ad5.GFP. $\Delta$ F, or other similarly constructed fiberless genome with either wild-type or modified fibers. Preferably for purposes herein the fibers are from an Ad serotype D virus, more preferably from Ad37.

#### EXAMPLE 6

##### 10           Pseudotyping and Infectivity of Recombinant Adenoviral Vectors               Produced with Improved Fiber-Complementing Cell Lines

##### A.       Pseudotyping of Ad5. $\beta$ gal. $\Delta$ F

To verify that adenoviral vectors were produced had altered tropisms, viral particles were purified from either 633 or 644 cells and were then Western  
15   blotted and probed with a polyclonal rabbit antibody against the Ad2 fiber (which detects the Ad5 and chimeric 5T3H fiber proteins).

B.       Infectivity of Cells with 633 or 644 Generated Virus Particles   The cell lines, 633 or 644, prepared as described above, were infected with the indicated number of particles/cell of Ad5. $\beta$ gal. $\Delta$ F and virus particles produced.  
20   Virus was then used to infect selected cell lines, including 211B, MRC-5 human fibroblasts, A-10 rat aortic endothelial cells, and THP-1 human monocytic cells. Unbound virus was removed by washing the cells and the cells were further incubated at 37°C for 48 hours. Cells were then fixed with glutaraldehyde and stained with X-gal. The percentage of stained cells was then determined by light  
25   microscopy where all experiments were done in triplicate.

The results indicated that adenoviral vectors could be retargeted by pseudotyping using packaging cell lines expressing different fiber proteins. Particles containing either fiber were equally infectious on 211B cells, while MRC-5 fibroblasts and THP-1 cells were more readily infected by virus containing  
30   the chimeric fiber. The A-10 rat endothelial cells were more readily infected by particles containing the wildtype Ad5 fiber protein.

-81-

## EXAMPLE 7

## Transient Transcomplementation

The ability of adenovirus type 5 (Ad5) to deliver therapeutic genes to cells is mediated by the interaction of the adenoviral fiber protein with the coxsackievirus-adenoviral receptor (CAR). Because a wide-range of cells express CAR, it was thought that it would be difficult to use adenoviruses to deliver genes to specific cell types. A system for testing modified fiber genes to identify tropisms of interest is described in copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000). An *in vitro* system has been developed that involves infection of tissue culture cells with a fiber-deleted Ad and transient co-transfection with a plasmid directing fiber expression. This system allows one to produce and evaluate modified fibers expressed on a viral particle. This system can be used to produce therapeutic quantities of adenoviral vectors with modified fiber proteins, with such fibers having a new tropism added by insertion of a desired ligand into the fiber gene. These fibers may also have the natural tropism (*i.e.* binding to CAR) ablated.

Plasmids used were pDV60 and pDV55 were prepared as described herein and in U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/US00/00265 on January 14, 2000). pDV60 is an pcDNA3.1-based expression plasmid that contains the CMV promoter, Ad5 tripartite leader, an intron, and the Ad5 fiber gene sequence. pDV55 contains no fiber gene and serves as the negative control. Ad5. $\beta$ gal. $\Delta$ F and 211B are described above. 293T cells are identical to 293 cells except they express an integrated SV40 large T antigen gene. HDF cells are human diploid fibroblasts. 293T cells express CAR and  $\alpha_v$  integrins; HDF cells express  $\alpha_v$  integrins but no CAR. Transfections with fiber expression plasmids were performed with Lipofectamine (GIBCO-BRL) using 20mg DNA and 50ml Lipofectamine per 15cm dish. Cells were maintained in DMEM supplemented with 10% fetal bovine serum.

The fiber deletion mutation of Ad5. $\beta$ gal. $\Delta$ F is complemented in *trans* by passaging virions through 211B, a cell line that stably expresses functional Ad5

-82-

fiber. The present system was designed to complement Ad5. $\beta$ gal. $\Delta$ F by modified fibers expressed from transfected episomal plasmids in 293T cells. The result is a simplified and rapid method to incorporate modified fibers on a viral particle containing the Ad5. $\beta$ gal. $\Delta$ F genome that does not require propagation of the virus.

5 The feasibility of transcomplementation of Ad5. $\beta$ gal. $\Delta$ F with episomal fiber-expressing plasmids was demonstrated in the following experiment. 293T cells were transfected with one of two plasmids: pDV55, which expresses no fiber or pDV60, which expresses wildtype Ad5 fiber. Fiber expression persists for at least six days. Twenty-four hours after transfection, these cells were infected at 2000 particles/cell with Ad5. $\beta$ gal. $\Delta$ F passaged through 211B cells. Seventy-two hours later, a crude viral lysate (CVL) was generated by exposing the cells to five freeze-thaw cycles. Viral particles were purified by cesium chloride gradient centrifugation. The resulting virions incorporated the fiber expressed from the episomal plasmid, as confirmed by Western blots performed with an antibody specific to the Ad5 fiber.

15 Episomal plasmid transcomplementation system is suitable for quickly expressing and evaluating the properties of modified fibers in the context of a viral particle. Episomal plasmid transcomplementation will also be of great utility for quickly evaluating a bank of modified fibers for other binding properties, including new tropisms and the ablation of the native tropism. In addition to the rapid generation and testing of large numbers of modified fibers, there are other advantages to the Ad5. $\beta$ gal. $\Delta$ F transcomplementation system in terms of production and safety. Episomal plasmid transcomplementation has the inherent advantage over transcomplementation in that it is not necessary to make a stable cell line for every modified fiber for complementation with Ad5. $\beta$ gal. $\Delta$ F. Because the Ad5. $\beta$ gal. $\Delta$ F is deleted in E1, E3 and fiber, there is an additional gene deletion, which should render it very suitable for gene therapy. In addition, the presence of the fiber gene deletion decreases the opportunity to generate replication-competent virus via recombination in the packaging cells. A single Ad vector preparation can be retargeted to any number of different cell types simply by transfecting the cells with the appropriate fiber-expression construct.

-83-

**EXAMPLE 8****Preparation of Adenoviral Gene Delivery Vectors Containing the Ad37 fiber protein**

This example describes construction of packaging cell lines expressing the Ad37 fiber protein, and their use in generating particles of a fiber-deleted Ad vector (such as Ad5. $\beta$ gal. $\Delta$ F) containing this fiber protein. The fiber protein is attached to the viral capsid by binding to the penton base protein through its N-terminus, and the Ad37 fiber was modified in order to make its N-terminal sequence more closely match that of the Ad5 protein to ensure that it would efficiently bind the Ad5 penton base in these vectors.

**A. Materials and methods**

Cell lines and wild-type adenovirus. Human A549 lung carcinoma epithelial cells and human Chang C conjunctival cells (American Type Culture Collection) were maintained in complete Dulbecco's Modified Eagle Medium (DMEM) with 10% fetal bovine serum. Wild-type Ad19p and Ad37 (ATCC) were propagated in A549 cells and purified by banding on CsCl<sub>2</sub> density gradients as previously described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Viral protein concentration was determined by the Bio-Rad Protein Assay, and was used to calculate the number of viral particles based on the known molecular weight of Ad2 virions ( $1 \mu\text{g} = 4 \times 10^9$  particles).

**B. Construction of the Ad37 fiber expressing cell lines and the recombinant Ad37 knob protein.****1. Construction of an Expression Plasmid for the Ad37 Fiber Protein (pDV80)**

The plasmid designated pDV80 (see, SEQ ID No. 42) prepared for expression of the Ad37 fiber protein in mammalian cells, uses the same regulatory elements as the elements in pDV60, pDV67, and pDV69 to express the Ad37 fiber in packaging lines. It was constructed in two steps.

First, the Ad37 fiber open reading frame was amplified from Ad37 genomic DNA using synthetic oligonucleotide primers, L37: 5' TGT CCT GGA TCC AAG ATG AAG CGC GCC CGC CCC AGC GAA GAT GAC TTC 3' (SEQ ID NO. 43) and 37FR: 5' AAA CAC GGC GGC CGC TCT TTC ATT CTT G 3' (SEQ ID NO. 44). L37 contains nucleotides (underlined) that differ from the Ad37

-84-

genomic sequence in order to add a unique *Bam*H1 site (bold) before the start codon (italicized) and to create point mutations that make the N-terminal sequence of the fiber more closely match the N-terminal sequence of the Ad5 fiber protein as follows:

5 Ad37 MSKRLRVEDDFNPVYPY (SEQ ID No. 45)  
           ↓ ↓ ↓ ↓ ↓  
           KRARPS (SEQ ID No. 46)  
 Ad5 MKRARPS EDTFNPVYPY (SEQ ID No. 47).

37FR also incorporates a unique *Not*1 site (bold). The PCR product was inserted  
 10 into the *Bam*H1 and *Not*1 sites of pCDNA3.1zeo(+) (Invitrogen) to create pDV78. The correct sequence of the Ad37 fiber protein, including inserted changes, was confirmed by sequencing.

Two point mutations in the fiber gene in the 705 line, S356 to P356 and I362 to T362, were discovered by the sequencing. The mutations are not in the  
 15 receptor binding domain in Ad37 fiber gene in the 705 cell line. They are buried in the knob trimer interface. To confirm that these mutations do not affect receptor binding, the Ad37 fiber protein with the correct sequence was recloned, and 293T cells transfected with the virus and subsequently infected with Ad5.GFPΔF to produce Ad37 pseudotyped virus. The results were the same as  
 20 the results of the experiments with Ad37 pseudotyped virus produced from line 705 (see, Wu *et al.* (2001) *Virology* 279:78-89).

Second, a 1.2 kb *Bam* H1/*Bgl* II fragment containing an adenovirus type 5 tripartite leader was excised from pDV55 (see EXAMPLE 3) and inserted into the *Bam* H1 site of pDV78 to create pDV80 (SEQ ID No 42). Plasmid pDV80 has  
 25 been deposited in the ATCC under accession number PTA-1147.

## 2. Construction of the recombinant Ad37 knob protein

Recombinant Ad37 knob protein containing an N-terminal T7•Tag was produced in *E. coli* using the PET expression system (Novagen). Ad37 fiber DNA (GenBank accession number U69132) was PCR amplified from wild-type Ad37  
 30 genomic DNA using the following primers (SEQ ID Nos. 48 and 49):  
 5' GGATCCATGGGATACTTGGTAGCA 3' (*Bam*HI site underlined and  
 5' GCAACTCGAGTCATTCTTGGGCAATATAGG 3' (*Xho*I site underlined).

-85-

The PCR reactions were performed at 94°C (denaturation), 55°C (annealing), 72°C (extension, 30 cycles) using *Taq* DNA polymerase (Qiagen). The amplified DNA fragments, which contained residues 172 to 365 of the Ad37 fiber protein with the addition of an N-terminal start codon (italicized), were purified and subcloned into the pCR-TOPO vector using the TA-Cloning Kit (Invitrogen). No replication errors were found by DNA sequencing. Plasmids from cultured transformed colonies were purified and digested with *Bam*HI and *Xho*I. The fragment was inserted into the *Bam*HI and *Xho*I sites of the bacterial expression vector, pET21a (Novagen), and transformed into (DE3)pLYS S expression cells (Invitrogen). Colonies were selected for knob expression by induction with 1 mM IPTG for four hours at 37°C and knob expression was determined by SDS-PAGE. The colony displaying highest knob expression was used for large-scale knob expression and induced with 0.5 mM IPTG at 30°C for four hours.

The recombinant T7•Tagged Ad37 knob protein was purified from sonicated bacteria using the T7•Tag Affinity Purification Kit as recommended by the manufacturer (Novagen). Recovered protein was analyzed for purity by SDS-PAGE followed by Coomassie staining or Western blotting with an HRP-conjugated  $\alpha$ -T7•Tag monoclonal antibody as described by the manufacturer (Novagen) or an  $\alpha$ -Ad37 fiber rabbit antibody.

### 3. Preparation of Cell Lines that Express the Ad37 fiber protein

Plasmid pDV80 DNA was purified using the Qiagen method and electroporated into the adenovirus-complementing cell line E1-2a S8 (see Examples herein; see also, Gorziglia *et al.* (1996) *J. Virology* 70:4173-4178; and Von Seggern *et al.* (1998) *J. Gen. Virol.* 79:1461-1468). Stable clones were selected with 600  $\mu$ g/ml zeocin (Invitrogen).

Clones were expanded and were screened for fiber expression by indirect immunofluorescence (Von Seggern *et al.* (1998) *J. Gen. Virol.* 79:1461-1468) using a rabbit polyclonal antibody directed against the Ad37 fiber ( $\alpha$ -Ad37 fiber rabbit antibody) raised by immunizing rabbits with recombinant Ad37 fiber protein. Two clones (lines 705 and 731) that expressed the protein at a uniformly high level were selected.

-86-

**EXAMPLE 9****Production of Pseudotyped Ad Vector Particles**

To generate vector particles equipped ('pseudotyped') with the Ad37 fiber protein, the Ad37 fiber-expressing 705 cells were infected (approximately  
5 1000 particles/cell) with Ad5. $\beta$ gal. $\Delta$ F or with Ad5.GFP. $\Delta$ F.

**Materials and methods****Ad5. $\beta$ gal. $\Delta$ F**

The construction of Ad5. $\beta$ gal. $\Delta$ F is described in Example 2 (it has been deposited on January 15, 1999, with the ATCC as listed above under accession  
10 number VR2636; see also, Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608; copending U.S. application Serial No. 09/482,682 filed January 14, 2000, and also International PCT application No. PCT/US00/00265, filed January 14, 2000).

**Ad5.GFP. $\Delta$ F**

15 Ad5.GFP. $\Delta$ F was constructed by recombination in bacteria using a modification of the AdEasy System (see, U.S. Patent No. 5,922,576; see, also He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514; the system is publicly available from the authors and other sources).

First, a fiber-deleted genomic plasmid was constructed by removing the  
20 fiber gene from pAdEasy-1 (see, U.S. Patent No. 5,922,576; and He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514; the AdEasy system and vectors are publicly available from He *et al.* at Johns Hopkins University). Plasmid pAdEasy-1 contains the entire Ad5 genome, except for nucleotides 1-3,533, which encompass the E1 genes, and nucleotides 28,,130-30,820, which  
25 encompass the E3 gene.

Plasmid pDV43 (see Example 2; see, also Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608) was digested with *Pac*1, the ends blunted by treatment with the large fragment of *E. coli* DNA polymerase and dNTPs, and the product re-ligated to produce plasmid pDV76. The resulting plasmid pDV76 is identical to  
30 pDV43 except for loss of the *Pac*1 site and contains the right end of the Ad5 genome with E3 and fiber deletions. A 4.23 kb fragment from pDV76 was amplified using the oligonucleotide primers (SEQ ID Nos. 50 and 51:



-87-

5' CGC GCT GAC TCT TAA GGA CTA GTT TC 3', including the unique *Spe*1 site in the Ad5 genome (bold); and 5' GCG CTT AAT TAA CAT CAT CAA TAA TAT ACC TTA TTT T 3', including a new *Pac*1 site (bold) adjacent to the right Ad5 ITR. Hence the resulting PCR amplified fragment contains nucleotides  
5 27,082 to 35,935 of the Ad5 genome with deletions of nucleotides 28,133 to 32,743 (the E3 and fiber genes), and was used to replace the corresponding *Spe*1/*Pac*1 fragment of pAdEasy 1 (see, U.S. Patent No. 5,922,576) to create pDV77.

Second, *E. coli* strain BJ5183 was electroporated with a mixture of  
10 pDV77 and *Pme*1-linearized pAdTrack as described (U.S. Patent No. 5,922,576; He *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2509-2514), and DNA was isolated from kanamycin-resistant colonies. The resulting plasmid, pDV83, contains a complete Ad5 genome with E1-, E3-, and fiber-deletions with a CMV-driven GFP reporter gene inserted at the site of the E1 deletion. The full length  
15 Ad chromosome was isolated by *Pac*1 digestion, and transfected into the E1- and fiber-complementing 633 cells (Von Seggern *et al.* (2000) *J. Virol.* 74:354-362). The 633 cells were produced by electroporating pDV67 (SEQ ID No. 30, deposited under ATCC accession number PTA-1145) into the E1-2a S8 cells, described above. The recovered virus Ad5.GFP.ΔF was then plaque purified by  
20 plating on 633 cells and virus stocks were prepared by freeze-thawing cell pellets.

#### Ad5-pseudotype particle production

##### Particles with Ad5 fiber

Ad5-pseudotyped particles were generated by virus growth in 633 cells,  
25 which express the wild type Ad5 fiber protein. Viral particles were isolated and purified over CsCl gradients (Von Seggern *et al.* (1999) *J. Virol.* 73:1601-1608; purified by centrifugation on preformed 15-40% CsCl gradients (111,000 x g for three hours at 4°C)). For analysis of viral proteins, ten μg of the purified particles were electrophoresed on 8-16% gradient gels and the protein  
30 transferred to nylon membranes. The resulting blot was probed with rabbit polyclonal antibodies raised against recombinant Ad37 fiber or Ad5 fiber or penton base proteins expressed in baculovirus-infected cells.

-88-

#### Particles with Ad37 fiber

Cells from the Ad37 fiber producing cell line 705 were infected at approximately 1000 particles/cell with Ad5. $\beta$ gal. $\Delta$ F or with Ad5.GFP. $\Delta$ F. Viral particles were isolated and purified over CsCl gradients. The bands were  
5 harvested, dialyzed into storage buffer (10 mM Tris-pH 8.1, 0.9% NaCl, and 10% glycerol), aliquoted and stored at -70°C.

#### Viral protein analyses

For analysis of viral proteins, 10  $\mu$ g of purified Ad5. $\beta$ gal. $\Delta$ F particles with no fiber (grown in 293 cells), the Ad5 fiber (grown in 633 cells), or the Ad37  
10 fiber (grown in 705 cells) were electrophoresed by 8-16% polyacrylamide gradient SDS-PAGE and the proteins were transferred to nylon membranes. The blot was then probed with  $\alpha$ -Ad37 fiber rabbit antibody. Ad5 fiber and penton base were detected by reprobing the blot with polyclonal antibodies raised against recombinant proteins expressed in baculovirus-infected cells (Wickman *et al.* (1993) *Cell* 2:309-319).  
15

#### Adenovirus infection and cell binding assays

Adherent Chang C and A549 cells were infected with GFP expressing Ad5 vectors containing the Ad5 fiber (Ad5.GFP. $\Delta$ F/5F) or the Ad37 fiber (Ad5.GFP. $\Delta$ F/37F) at 10,000 particles per cell for 3 hours at 37°C, 5% CO<sub>2</sub> in  
20 DMEM, 10% FCS. Cells were washed twice with saline and then cultured overnight at 37°C, 5% CO<sub>2</sub>. The next day, the cells were detached with buffer containing 0.05% (w/v) trypsin and 0.5 mM EDTA (Boehringer Mannheim) for 5 minutes at 37°C. Suspended cells were washed once with PBS and then resuspended in phosphate-buffered saline (PBS), pH 7.4. GFP fluorescence was  
25 measured with a FACScan flow cytometer. A threshold established by the fluorescence of uninfected cells was used to distinguish cells expressing GFP. To assess the role of CAR in Ad infection, 10,000 attached cells were pre-incubated with 180  $\mu$ g/ml RmcB, a function-blocking anti-CAR monoclonal antibody (Hsu *et al.* (1988) *J. Virol.* 62:1647-1652), in complete DMEM for 1  
30 hour at 4°C. A small volume containing Ad5.GFP. $\Delta$ F/5F or Ad5.GFP. $\Delta$ F/37F was then added at 10,000 particles per cell. The cells were infected for 3 hours, cultured overnight, harvested, and analyzed for GFP expression. Percent

-89-

cells expressing GFP was determined by the percent of cells detected above a threshold set by the fluorescence of uninfected Chang C cells.

To measure adenovirus binding to cells, wild type Ad37 was labeled with  $^{125}\text{I}$  using Iodogen (Pierce) according to manufacturer instructions and separated  
5 from free  $^{125}\text{I}$  by gel filtration as described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Binding of radiolabeled wild type Ad37 on Chang C cells was then quantitated as described (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Non-specific binding was determined by incubating cells and labeled Ad37 particles in the presence of 100-fold concentration of unlabeled Ad37. Specific  
10 binding was calculated by subtracting the non-specific binding from the total cpm bound. To examine if divalent cations are required for binding, 10 mM ethylenediaminetetraacetic acid (EDTA) or various concentrations of  $\text{CaCl}_2$ , or  $\text{MgCl}_2$  were added to cells before incubation with labeled virus. To examine if the receptor for Ad37 is a protein, cells were pretreated with 10  $\mu\text{g/ml}$  trypsin  
15 (GIBCO), subtilisin (Sigma), proteinase K (Boehringer-Mannheim), and bromelain (Sigman) at 37 °C for 1 hour, then washed twice with complete DMEM before adding labeled virus. Cells were >95% viable after protease treatment.

Ad37 binding to conjunctival cells is calcium-dependent. Specific  $^{125}\text{I}$ -labeled Ad37 binding to Chang C cells was measured in the presence of 10 mM  
20 EDTA and in the presence of varying concentrations of calcium chloride or magnesium chloride. Specific binding was determined by subtracting the nonspecific counts in the presence of 100-fold excess unlabeled virus from the total counts.

Pretreatment of conjunctival cells with proteases inhibits Ad37 binding.  
25 Change C cells were pretreated with various proteases for 1 hour before binding  $^{125}\text{I}$ -labeled Ad37 to the cells. Nonspecific binding was measured by adding 100-fold unlabeled Ad37 to cells with  $^{125}\text{I}$ -labeled Ad37 and subtracting from total counts for specific binding. Percent inhibition represents the difference in specific binding of untreated cells and pretreated cells as a percentage of the  
30 specific binding of untreated cells.

-90-

**Virus overlay protein blot assay (VOPBA)**

For VOPBA of human conjunctival membrane proteins probed with Ad37 in the presence of EDTA or calcium chloride, Chang C membrane fractions were separated by 8% SDS-PAGE and transferred to a PDVF membrane. The

5 membrane was subsequently probed with or without whole Ad37 particles, a polyclonal antibody against Ad37 fiber, and finally a horseradish peroxidase conjugated anti-rabbit antibody, in the presence of EDTA or calcium chloride. Transferred Chang C membrane proteins were probed with recombinant Ad37 knob protein, instead of Ad37 knob, in the presence of calcium chloride.

10 Confluent monolayers of Chang C and A549 cells were detached by scraping, pelleted by centrifugation, and then resuspended in 250 mM sucrose, 20 mM HEPES, pH 7.0, 1 mM EDTA, and 2  $\mu$ g/ml aprotinin and leupeptin. Cells were transferred into a dounce homogenizer and disrupted with 30 strokes. Organelles and nuclei were pelleted at 500g for 15 min. Plasma membrane  
15 fragments were then pelleted from the supernatant of cell lysates at 200,000g for 1 hour and then resuspended in 10 mM Tris•Cl, pH 8.1, 10  $\mu$ g/ml aprotinin and leupeptin.

Cell membranes of Chang C or A549 cells were incubated (1:1) with a  
2% SDS, non-reducing buffer and separated on an 8% polyacrylamide gel  
20 without boiling. Membrane proteins were then electroblotted onto a PVDF membrane (Immobilon-P) and blocked in 5% (w/v) milk in PBS, pH 7.4, 0.02% Tween-20 (PBS-T). After blocking, the membrane was incubated with 1  $\mu$ g/ml wild-type Ad19p or Ad37 in 0.5% (w/v) milk in PBS-T, 1 mM  $\text{CaCl}_2$ , for 1 hour at room temperature. The membrane was then washed once with phosphate-  
25 buffered saline, pH 7.4 (PBS), 1 mM  $\text{CaCl}_2$ , and incubated with 1:500 dilution of  $\alpha$ -Ad37 fiber rabbit antibody in 0.5% (w/v) milk in PBS-T, 1 mM  $\text{CaCl}_2$ , for 30 minutes at room temperature. The membrane was washed again with PBS, 1 mM  $\text{CaCl}_2$ , and incubated with 1:5000 dilution of horseradish peroxidase (HRP) conjugated  $\alpha$ -rabbit antibody (Sigma) in 0.5% (w/v) milk in PBS-T, 1 mM  $\text{CaCl}_2$ ,  
30 for 30 minutes at room temperature. The membrane was washed four times in PBS, 1 mM  $\text{CaCl}_2$ , once with PBS-T, 1 mM  $\text{CaCl}_2$ , and once in 1 mM  $\text{CaCl}_2$ . The blot was developed with enhanced chemiluminescence reagents (Pierce) for 5

-91-

minutes and placed onto a piece of Biomax film (Kodak) for 5 seconds to 1 minute. For divalent metal cation experiments, membranes were incubated in the presence of 2 mM EDTA instead of 1 mM  $\text{CaCl}_2$  in all solutions. To assay fiber knob binding to cell membrane proteins, membrane filters were incubated with 1  $\mu\text{g/ml}$  purified T7-tagged Ad37 knob protein in Tris-buffered saline, 0.1% Tween-20, 1 mM  $\text{CaCl}_2$ , for 1 hour at room temperature.  $\alpha$ -Ad37 fiber rabbit antibody and HRP-conjugated anti-rabbit antibody were applied and the membrane was developed with substrate solution as described above.

**Results: Comparison of adenovirus infection of human conjunctival and lung epithelial cells with virus particles retargeted with Ad5 or Ad37 fiber proteins**

Packaging cell lines producing the Ad37 fiber protein were generated. Since the N-terminal amino acid sequences of the Ad5 and Ad37 fiber proteins differ significantly, and to ensure that the Ad37 fiber would be efficiently incorporated into Ad5 vector particles, several residues in the wild-type Ad37 fiber were mutated to more closely match the Ad5 sequence. Stable cell lines producing this fiber under control of the CMV promoter and the adenovirus type 5 tripartite leader were then generated and screened for fiber expression by indirect immunofluorescence. One clone (line 705), which expressed the Ad37 fiber at a high level, was selected for further study.

Cells from one cell line 633, which expresses the wild-type Ad5 fiber protein, and line 705 were infected with a fiber-deleted Ad5 vector carrying a  $\beta$ galactosidase reporter gene. The resulting vector particles contained the Ad5 fiber protein (Ad5. $\beta$ gal. $\Delta$ F/5F) and the Ad37 fiber protein (Ad5. $\beta$ gal. $\Delta$ F/37F), respectively. Incorporation of the correct fiber protein into viral particles was verified by Western blotting. Adenoviral vectors containing the GFP reporter gene, Ad5.GFP. $\Delta$ F/5F and Ad5.GFP. $\Delta$ F/37F, were created in the same fashion.

Infection of a variety of cell types using the retargeted adenovirus particles was examined. As assayed by GFP fluorescence, Ad5.GFP. $\Delta$ F/5F exhibited good gene delivery to lung epithelial (A549) and conjunctival cells (Chang C). In contrast, Ad5.GFP. $\Delta$ F/37F efficiently delivered GFP to Chang C cells, but exhibited very poor gene delivery to A549 cells. Although CAR is expressed on the surface of A549 cells, as indicated by AD5.GFP. $\Delta$ F/5F

-92-

infection, Ad5.GFP. $\Delta$ F/37F was unable to infect these cells efficiently. This experiment shows that the Ad37 fiber protein can confer preferential infection of human conjunctival cells, but not CAR-expressing human lung epithelial cells.

Hence CAR is not the primary receptor for Ad37. Recent studies reported  
5 that expression of CAR on the surface of chinese hamster ovary (CHO) cells did not improve Ad37 binding (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), implying that Ad37 does not use CAR as a primary receptor. In order to verify this on human conjunctival cells, A549 and Chang C cells were pretreated with RmcB (Hsu *et al.* (1988) *J. Virol.* 62:1647-1652), a function-blocking monoclonal  
10 antibody against CAR. The RmcB antibody inhibited infection of A549 cells by Ad5.GFP. $\Delta$ F/5F, but it had little effect on infection of Chang C cells by Ad5.GFP. $\Delta$ F/37F. This indicates that CAR is not the primary receptor for Ad37 on Chang C conjunctival cells.

Ad37 binding to conjunctival cells requires divalent metal cations. It has  
15 been proposed (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915) that a combination of fiber binding to CAR and penton base binding to  $\alpha_v$ -integrins allows some adenovirus serotypes to attach to cells. Although  $\alpha_v$ -integrin binding to the RGD motif of the adenovirus penton base is of relatively low affinity (Wickman *et al.* (1993) *Cell* 2:309-319), it may nonetheless contribute to viral attachment to the  
20 cell surface. Ad37 shows a particularly strong affinity for binding to integrin  $\alpha_v\beta_5$  (Mathias *et al.* (1998) *J. Virol.* 72:8669-8675), suggesting that integrin  $\alpha_v\beta_5$  might be a primary receptor for Ad37. Binding of the RGD motif by  $\alpha_v$ -integrins requires the presence of divalent cations, such as calcium or magnesium (Stuiver *et al.* (1996) *J. Cell Physiol.* 168:521-531). In contrast, no divalent cations were  
25 required for binding in the CAR-Ad12 knob complex (Bewley *et al.* (1999) *Science* 286:1579-1583).

To investigate the potential role of  $\alpha_v$ -integrins and divalent metal cations in Ad37 receptor binding,  $^{125}$ I-labeled Ad37 binding to Chang C cells was examined in the absence or presence of EDTA. EDTA inhibited Ad37 binding to  
30 conjunctival cells but did not alter Ad5 binding. These findings suggest a requirement for divalent metals for Ad37 binding.

-93-

The presence of either calcium or magnesium ions helps  $\alpha_v\beta_5$  organize in focal contacts (Stuiver *et al.* (1996) *J. Cell Physiol.* 168:521-531), suggesting that calcium and magnesium aid in integrin  $\alpha_v\beta_5$  function. To further test the potential role of integrin  $\alpha_v\beta_5$  in Ad37 cell attachment,  $^{125}\text{I}$ -labeled Ad37 binding to Chang C cells was measured in the presence of varying concentrations of calcium or magnesium chloride. Magnesium ions had little effect on Ad37 binding to Chang C cells. In contrast, calcium ions dramatically enhanced Ad37 binding to Chang C cells. The optimal concentration of calcium chloride for Ad37 binding was 1 mM, while higher concentrations of calcium actually decreased virus binding to cells. The fact that calcium, but not magnesium, promoted Ad37 attachment is not consistent with integrin  $\alpha_v\beta_5$  as the primary receptor for viral attachment to the cells since either metal will support ligand binding to integrin  $\alpha_v\beta_5$ . Moreover, A549 cells express abundant  $\alpha_v$ -integrins (Mathias *et al.* (1998) *J. Virol.* 72:8669-8675) but were unable to support efficient binding of Ad37.

Wild-type Ad37 particles bind to three conjunctival membrane proteins. Recent studies reported that protease treatment of CHO cells abolished Ad37 binding (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), implying that Ad37 bound to a protein receptor on CHO cells. Scatchard analysis of Ad37 binding to Chang C cells showed that each cell expresses approximately 24,000 fiber binding sites (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). To determine if the Ad37 binding site on human conjunctival cells is also a protein, Chang C cells were treated with different proteases prior to measuring binding of  $^{125}\text{I}$ -labeled Ad37. Digestion of surface proteins by all four proteases inhibited Ad37 binding to Chang C cells by greater than 50%. This finding showed that Ad37 also binds to a protein receptor on Chang C cells.

Virus overlay protein blot assays (VOPBAs) were used to identify candidate viral protein receptors. This Western blot technique uses intact viral particles in place of antibodies to probe viral-receptors interactions. VOPBA was used herein to identify Chang C membrane proteins that bind to Ad37. In the absence of Ad37 particles, no protein bands were observed, while addition of virus in the absence of calcium revealed binding to a single 45 kDa protein. In

-94-

the presence of 1 mM calcium chloride, Ad37 reacted with three proteins with approximate molecular weights of 45, 50 and 60 kDa. The same three proteins were detected using a recombinant Ad37 fiber knob alone, indicating that Ad37-receptor interactions are fiber mediated and do not involve interactions of other capsid proteins such as the penton base. The size of the calcium-independent protein (45 kDa) is very similar to the known molecular weight of CAR. A direct comparison of the Ad37 VOPBA and a CAR Western blot showed that the 45 kDa receptor co-migrates with CAR on SDS-PAGE. Moreover, two other members of subgroup D adenoviruses, Ad9 and AD15, have been shown to bind to CAR (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915).

Since CAR does not appear to mediate Ad37 binding on intact Chang C cells, the possibility that the 50 or 60 kDa protein serves this function was tested by examining an adenovirus serotype that does not bind to Chang C cells. Ad19p, a closely related subgroup D adenovirus, binds poorly to Chang C cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802) and Ad19p recognition of the Ad37 receptor is therefore unlikely. Ad19p particles bound to the 45 and 60 kDa receptors in the VOPBA, but did not bind to the 50 kDa receptor. Moreover, the 50 kDa receptor is expressed on Chang C cells, but not A549 cells, which only support low levels of Ad37 binding and infection. Taken together, these data indicate that the 50 kDa protein is a primary candidate receptor for Ad37 on human conjunctival cells.

#### Discussion

The identification of the CAR protein as a major adenovirus receptor does not explain why certain subgroup D members, such as Ad37, preferentially infect ocular cells. A 50 kDa human conjunctival cell membrane protein is identified herein as a primary candidate for the receptor for Ad37. This 50 kDa protein is not present on A549 lung epithelial cells. Ad37 binding to this receptor is calcium-dependent, which is consistent with Ad37 binding and infection experiments. Ad37 also bound to a 60 kDa protein that is present on human conjunctival and lung epithelial cells. It does not, however, appear to be serotype specific. The molecular weights of MHC class I heavy chain, which has been proposed as a receptor for Ad5, and  $\alpha\beta_3$  and  $\alpha\beta_5$  integrins, receptors for



-95-

the penton base, are distinct from the 50 or to kDa receptor characterized in this study.

The studies of Ad37-receptor interaction using VOPBAs are consistent with previous studies showing that subgroup D adenoviruses can bind to the extracellular domain of CAR (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915). Biochemical and structural studies on knob-CAR interactions indicate that the CAR binding site is located on the AB-loop of the fiber knob. Alignment of the fiber sequences of Ad37 and other adenoviruses reveals that the AB-loop of Ad37 is similar to those of Ad12 and Ad5. Moreover, a phylogenetic tree of adenovirus knobs (Roelvink *et al.* (1998) *J. Virol.* 72:7909-7915) shows that fiber proteins of subgroup D are similar to those of subgroup C and E, which use CAR as their primary receptor. Ad37 does not, however, appear to effectively use CAR as a primary receptor, as demonstrated by virus binding and infection studies on Chang C conjunctival cells and A549 lung epithelial cells.

It has been reported that Ad37 uses sialic acid as a receptor on chinese hamster ovary (CHO) cells and human lung carcinoma (A549) cells (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). Human conjunctival cells were not studied. Human corneal epithelial (HCE) cells were the only ocular cell line studied and Ad37 binds relatively poorly to these cells, compared to binding on A549 cells (Arnberg *et al.* ((2000) *J. Virol.* 74:42-48). In addition,  $8,4 \times 10^7$  wheat germ agglutinin molecules per cell were required to significantly inhibit Ad37 binding to sialic acid on sialic acid positive CHO cells (Arnberg *et al.* (2000) *J. Virol.* 74:42-48), three orders of magnitude higher than the number of Ad37 receptors on Chang C conjunctival cells (Huang *et al.* (1999) *J. Virol.* 73:2798-2802). Clearly, sialic acid is not the only factor responsible for Ad37 binding to the cell surface and its influence on Ad37 tropism is unclear.

The results herein show that Ad37 selects a 50 kDa cellular receptor for binding to conjunctival cells, but it is possible that sialic acid also plays a role in this interaction. The characterization and identification of the Ad37 receptor have therapeutic implications and also explain the different tropism of Ad37. The 50 kDa receptor for Ad37 may also be the receptor for other subgroup D adenoviruses that cause severe cases of EKC, Ad19a and Ad8. Ad19p is a

-96-

nonpathogenic variant of Ad19 (Amberg *et al.* (1998) *Virology* 227:239-244) while Ad19A, along with Ad8 and Ad37, are major causes of EKC. Ad19a and Ad37 have identical fiber proteins (Arnberg *et al.* (1998) *Virology* 227:239-244) and have similar tropism in vivo. Ad8, Ad19a, and Ad37 agglutinate dog and guinea pig erythrocytes more effectively than four other serotypes that are associated with less severe forms of conjunctivitis (Arnberg *et al.* (1998) *Virology* 227:239-244), implying that the receptors of Ad18, Ad19A, and Ad37 have similar characteristics. Hence, this 50 kDa receptor is an attractive drug target against EKC caused by adenoviruses to provide therapeutic intervention of ocular diseases associated with these viruses.

#### EXAMPLE 10

##### Targeting of the Ad5 vector to photoreceptor cells

The fiber-deleted adenovirus vector Ad5.GFP. $\Delta$ F was propagated in 705 cells, which express a modified Ad37 fiber protein. Viral particles (Ad5.GFP. $\Delta$ f/37F) were harvested, CsCl-purified and dialized into 0.9% NaCl, 10 mM Tris, pH 8.1, and 10% glycerol. Two to three  $\mu$ l of the resulting solution, containing approximately  $1 \times 10^9$  particles/ $\mu$ l was injected into the vitreous chamber of a mouse eye. Seven days post-injection, eyes were harvested, fixed with paraformaldehyde and cryo-sectioned. Sections were stained with an anti-rhodopsin antibody to identify photoreceptor cells and with DAPI to show all cell nuclei. The resulting sections showed red anti-rhodopsin staining in the photoreceptors, blue DAPI-stained nuclei, and green GFP staining in any transduced cells. The results revealed substantially exclusive transduction of photoreceptors. Co-localization of rhodopsin staining and GFP expression indicated selective transduction of photoreceptor cells.

As a control, contralateral eyes were injected with a stock of the fiber-deleted vector AD5. $\beta$ gal. $\Delta$ F grown in the same Ad37 fiber-expressing cells. Since this virus (Ad5. $\beta$ gal. $\Delta$ F/37F) produces  $\beta$ gal rather than GFP, the green staining is absent from the photoreceptors.

Additional experiments using the AD37 fiber for targeting to the photoreceptor cells have been performed. Subretinal and intravitreal injection have been used in mouse models and the results demonstrate targeting to the

-97-

photoreceptors. As with intravitreally injected eyes, the major cell type infected via subretinal administration was the photoreceptor.

As noted, Ad5.GFP. $\Delta$ F /37F infected Chang C cells efficiently, but A549 cells poorly. Ad37 fiber protein confers preferential infection on human  
5 conjunctival cells, but not CAR-expressing human lung epithelial cells. Binding to conjunctival cells requires divalent cations.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

-98-

**WHAT IS CLAIMED IS:**

1. An isolated nucleic acid molecule, comprising:  
adenovirus inverted terminal repeat sequences; an adenovirus packaging  
signal operatively linked thereto; and a photoreceptor-specific promoter.
- 5 2. The isolated nucleic acid molecule of claim 1, further comprising a  
nucleic acid encoding a therapeutic product operatively linked to the promoter.
3. The isolated nucleic acid molecule of claim 1, wherein the  
promoter is a rhodopsin promoter.
4. The nucleic acid molecule of claim 1, wherein the adenovirus  
10 genome does not encode a functional fiber protein such that packaging the  
nucleic acid requires complementation in a packaging cell.
5. A recombinant adenovirus vector, comprising the nucleic acid  
molecule of any of claims 1-4 packaged therein.
6. A recombinant adenovirus vector of claim 5, wherein inverted  
15 terminal repeat sequences (ITR) and a packaging signal are derived from  
adenovirus type 2 or adenovirus type 5.
7. A recombinant adenovirus vector of claim 5, wherein the virus  
comprises a fiber protein.
8. A recombinant adenovirus vector of claim 7, wherein the fiber  
20 protein selectively binds to photoreceptors in the eye of a mammal.
9. A recombinant adenovirus vector of claim 7, wherein the fiber is a  
chimera composed of N-terminal sequences from adenovirus type 2 or type 5,  
and a sufficient portion of an adenovirus serotype D fiber for selective binding to  
photoreceptors in the eye of a mammal.
- 25 10. A method for targeted delivery of a gene product to the eye of a  
mammal, comprising:  
administering a recombinant adenovirus virus that comprises heterologous  
DNA encoding the gene product or resulting in expression of the gene product,  
wherein the recombinant virus comprises a fiber protein that specifically or  
30 selectively binds to receptors that are expressed on cells in the eye.

-99-

11. The method of claim 10, wherein the cells are photoreceptors.
12. The method of claim 10, wherein administration is effected by intraocular delivery.
13. The method of claim 10, wherein administration is effected by a  
5 method selected from subretinal injection, intravenous administration, periorbital administration, and intravitreal administration.
14. The method of claim 10, wherein the recombinant virus comprises a fiber protein from an adenovirus type D serotype.
15. The method of any of claims 10-14, wherein the fiber protein is an  
10 adenovirus type 37.
16. The method of any of claims 10-14, wherein the fiber is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fiber protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient portion of an adenovirus serotype D knob portion of the  
15 fiber for selective binding to photoreceptors in the eye of a mammal.
17. The method of any of claims 10-16, wherein the recombinant virus is an adenovirus type D serotype.
18. The method of any of claims 10-17, wherein the encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a  
20 nucleic acid comprising the therapeutic product.
19. The method of claim 18, wherein the therapeutic product is selected from the group consisting of a trophic factor, an anti-apoptotic factor, a gene encoding a rhodopsin protein, a wild-type Stargardt disease gene (STDG1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-  
25 specific gene product.
20. The method of any of claims 10-19, wherein delivery is effected for treatment of an ocular disease.
21. The method of claim 20, wherein the disorder is a retinal degenerative disease.
- 30 22. The method of claim 20, wherein the disease is retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularization, or retinoblastoma.

-100-

23. The method of any of claims 10-22, wherein the mammal is a human.

24. The method of any of claims 10-22, wherein the viral nucleic acid comprises:

- 5        an adenovirus inverted terminal repeat (ITR) sequences; and  
      an adenovirus packaging signal operatively linked thereto.

25. The method of claim 24, wherein the ITRs and packaging signal are derived from an adenovirus serotype B or C.

26. The method of claim 24, wherein the ITRs and packaging signal  
10 are derived from an adenovirus type 2 or 5.

27. The method of claim 24, wherein the viral nucleic acid further comprises a photoreceptor-specific promoter.

28. A method of targeted gene therapy, comprising:  
      administering a recombinant viral vector that comprises an adenovirus  
15 type 37 fiber protein or portion thereof, whereby the vector selectively transduces photoreceptors and delivers a gene product encoded by the recombinant viral vector; wherein the portion is sufficient for selective binding to photoreceptors.

29. The method of claim 28, wherein the vector is administered into  
20 the eye.

30. The method of claim 28, wherein the vector is administered to the vitreous cavity of the eye.

31. The method of claim 28, wherein administration is effected by subretinal injection, intravenous administration, periorbital administration or  
25 intravitreal administration.

32. The method of any of claims 10-31, wherein at least about  $10^7$  plaque forming units of virus are administered.

33. The method of any of claims 10-31, wherein about 1 plaque forming unit to about  $10^{14}$  plaque forming units of virus are administered.

## SEQUENCE LISTING

<110> VON SEGGERN, DANIEL  
NEMEROW, GLEN R.  
FRIEDLANDER, MARTIN

<120> VECTORS FOR OCULAR TRANSDUCTION AND USE THEREFOR FOR GENETIC THERAPY

<130> 756.1PCT/NOV0205P

<140>

<141> 2001-05-01

<150> 09/562,934

<151> 2000-05-01

<160> 51

<170> PatentIn Ver. 2.1

<210> 1

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 1

cggtacacag aattcaggag acacaactcc

30

<210> 2

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 2

gcctggatcc gggaagttac gtaacgtggg aaaac

35

<210> 3

<211> 12

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: linker

<400> 3

cgcggtatccg cg

12

<210> 4

<211> 8710

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Description of Artificial Sequence: plasmid

&lt;400&gt; 4

|            |             |             |             |             |             |      |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| cacctaaatt | gtaagcggtta | atattttgtt  | aaaatttcg   | ttaaattttt  | gttaaatcag  | 60   |
| ctcatttttt | aaccaatagg  | ccgaaatcgg  | caaaatccct  | tataaatcaa  | aagaatagac  | 120  |
| cgagataggg | ttgagtgttg  | ttccagtttg  | gaacaagagt  | ccactattaa  | agaacgtgga  | 180  |
| ctccaacgtc | aaagggcgaa  | aaaccgtcta  | tcagggcgat  | ggcccactac  | gtgaaccatc  | 240  |
| accctaatac | agttttttgg  | ggctcaggtg  | ccgtaaagca  | ctaaatcgga  | accctaagg   | 300  |
| gagcccccca | tttagagctt  | gacggggaaa  | gccggcgaa   | gtggcgagaa  | aggaagggaa  | 360  |
| gaaagcgaaa | ggagcgggcg  | ctagggcgct  | ggcaagtgt   | gcggtcacgc  | tgcgcgtaac  | 420  |
| caccacaccc | gccgcgctta  | atgcgccgct  | acagggcgcg  | tcccattcgc  | cattcaggct  | 480  |
| gcgcaactgt | tgggaagggc  | gatcgggtcg  | ggcctcttcg  | ctattacgcc  | agctggcgaa  | 540  |
| agggggatgt | gctgcaaggc  | gattaagtgt  | ggtaacgcca  | gggttttccc  | agtcacgacg  | 600  |
| ttgtaaaacg | acggccagtg  | aattgttaata | cgactcacta  | tagggcggaat | tgggtaccgg  | 660  |
| gccccccctc | gaggtcgacg  | gtatcgataa  | gcttgatata  | gaattcagga  | gacacaactc  | 720  |
| caagtgcata | ctctatgtca  | ttttcatggg  | actgggtctg  | ccacaactac  | attaatgaaa  | 780  |
| tatttgccac | atcctcttac  | actttttcat  | acattgcccc  | agaataaaga  | atcgtttgtg  | 840  |
| ttatgttttc | acgtgtttat  | ttttcaattg  | cagaaaattt  | caagtcattt  | ttcattcagt  | 900  |
| agtatagccc | caccaccaca  | tagcttatac  | agatcacctg  | accttaatac  | aactcacaga  | 960  |
| accctagtat | tcaacctgcc  | acctccctcc  | caacacacag  | agtacacagt  | cctttctccc  | 1020 |
| cggctggcct | taaaaagcat  | catatcatgg  | gtaacagaca  | tattcttagg  | tgttatattc  | 1080 |
| cacacgggtt | cctgtcgagc  | caaacgctca  | tcagtgatat  | taataaactc  | cccgggcagc  | 1140 |
| tcacttaagt | tcattgtcgt  | gtccagctgc  | tgagccacag  | gctgctgtcc  | aacttgccgt  | 1200 |
| tgtttaacgg | gcggcggaag  | agaagtccac  | gcctacatgg  | gggtagagtc  | ataatcgtgc  | 1260 |
| atcaggatag | ggcggtgggt  | ctgcagcagc  | gcgcgaataa  | actgctgccc  | ccgcccgtcc  | 1320 |
| gtcctgcagg | aatacaacat  | ggcagtggtc  | tcctcagcga  | tgattogcac  | cgcccgcagc  | 1380 |
| ataaggcgcc | ttgtcctccg  | ggcacagcag  | cgcaccctga  | tctcacttaa  | atcagcacag  | 1440 |
| taactgcagc | acagcaccac  | aatattgttc  | aaaatccccc  | agtgcagggc  | cgtgtatcca  | 1500 |
| aagctcatgg | cggggaccac  | agaacccacg  | tgcccatcat  | accacaagcg  | caggtagatt  | 1560 |
| aagtggcgac | ccctcataaa  | cacgctggac  | ataaacatta  | cctcttttgg  | catgttgtaa  | 1620 |
| ttcaccacct | cccgggtacca | tataaacctc  | tgattaaaca  | tggcgccatc  | caccaccatc  | 1680 |
| ctaaaccagc | tggccaaaac  | ctgcccgcgc  | gctatacact  | gcaggggaac  | gggactggaa  | 1740 |
| caatgacagt | ggagagccca  | ggactcgtaa  | ccatggatca  | tcattgctcg  | catgatatac  | 1800 |
| atgttggcac | aacacaggca  | cacgtgcata  | cacttcctca  | ggattacaag  | ctcctcccgc  | 1860 |
| gttagaacca | tatcccaggg  | aacaacccat  | tcctgaatca  | gcgtaaatcc  | cacactgcag  | 1920 |
| ggaagacctc | cacgtaactc  | cacgttgtgc  | attgtcaaag  | tgttacattc  | gggcagcagc  | 1980 |
| ggatgatcct | ccagtatggg  | agcgcgggtt  | tctgtctcaa  | aaggaggtag  | acgatcccta  | 2040 |
| ctgtacggag | tgccgcgaga  | caaccgagat  | cgtgttggtc  | gtagtgtcat  | gccaaatgga  | 2100 |
| acgcgggacg | tagtcatatt  | tcctgaagca  | aaaccagggt  | cgggcgtgac  | aaacagatct  | 2160 |
| gcgtctccgc | tctcgccgct  | tagatcgctc  | tggtgtagtag | ttgtagtata  | tccactctct  | 2220 |
| caaagcatcc | aggcgcccc   | tggtctcggg  | ttctatgtaa  | actccttcat  | gcgcccgtgc  | 2280 |
| cctgataaca | tccaccaccg  | cagaataagc  | cacaccacgc  | caacctacac  | attcgttctg  | 2340 |
| cgagtccac  | cggggaggag  | cgggaagagc  | tggaagaacc  | atgttttttt  | ttttattcca  | 2400 |
| aaagattatc | caaaacctca  | aaatgaagat  | ctattaagtg  | aacgcgctcc  | cctccggtgg  | 2460 |
| cgtgggtcaa | ctctacagcc  | aaagaacaga  | taatggcatt  | tgtaagatgt  | tgcacaaatg  | 2520 |
| cttccaaaag | gcaaacggcc  | ctcacgtcca  | agtggacgta  | aaggctaaac  | ccttcagggt  | 2580 |
| gaatctcctc | tataaacatt  | ccagcacctt  | caaccatgcc  | caaataattc  | tcattctcgc  | 2640 |
| accttctcaa | tatatctcta  | agcaaatccc  | gaatattaag  | tcgggccatt  | gtaaaaatct  | 2700 |
| gctccagagc | gccctccacc  | ttcagcctca  | agcagcgaat  | catgattgca  | aaaattcagg  | 2760 |
| ttcctcacag | acctgtataa  | gattcaaaag  | cggaaacatta | acaaaaatac  | cgcatcccg   | 2820 |
| taggtccctt | cgcagggcca  | gctgaacata  | atcgtgcagg  | tctgcacgga  | ccagcgccgc  | 2880 |
| cacttccccg | ccaggaacct  | tgacaaaaga  | acccacactg  | attatgacac  | gcatactcgg  | 2940 |
| agctatgcta | accagcgtag  | ccccgatgta  | agctttgttg  | catgggcccg  | gatataaaat  | 3000 |
| gcaaggtgct | gctcaaaaaa  | tcaggcaaa   | cctcgcgcaa  | aaaagaaaag  | acatcgtagt  | 3060 |
| catgtctcat | cagataaagg  | caggttaagc  | ccggaaccac  | cacagaaaaa  | gacaccattt  | 3120 |
| ttctctcaaa | catgtctcgc  | ggtttctgca  | taaacacaaa  | ataaaaatac  | aaaaaaacat  | 3180 |
| ttaaacatta | gaagcctgtc  | ttacaacagg  | aaaaacaacc  | cttataagca  | taagacggac  | 3240 |
| tacggccatg | cggcggtgac  | cgtaaaaaaa  | ctggtcaccg  | tgattaaaaa  | gcaccaccga  | 3300 |
| cagctcctcg | gtcatgtccg  | gagtcataat  | gtaagactcg  | gtaaacacat  | cagggttgatt | 3360 |
| catcggtcag | tgctaaaaag  | cgaccgaaat  | agcccggggg  | aatacatacc  | cgcaggcgta  | 3420 |
| gagacaacat | tacagcccc   | ataggaggta  | taacaaaatt  | aataggagag  | aaaaaacacat | 3480 |
| aaacacctga | aaaaccctcc  | tgcttaggca  | aaatagcacc  | ctcccgcctc  | agaacaacat  | 3540 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| acagcgcttc  | acagcggcag  | cctaacagtc  | agccttacca  | gtaaaaaaga  | aaacctatta  | 3600 |
| aaaaaacacc  | actcgacacg  | gcaccagctc  | aatcagtcac  | agtgtaaaaa  | agggccaagt  | 3660 |
| gcagagcgag  | tatatatagg  | actaaaaaat  | gacgtaacgg  | ttaaagtcca  | caaaaaacac  | 3720 |
| ccagaaaacc  | gcacgcgaac  | ctacgcccag  | aaacgaaagc  | caaaaaaccc  | acaacttcct  | 3780 |
| caaatcgta   | cttccggttt  | cccacgttac  | gtaacttccc  | ggatccgcgg  | cattcacagt  | 3840 |
| tctccgcaag  | aattgattgg  | ctccaattct  | tggagtgggtg | aatccgtag   | cgaggtgccg  | 3900 |
| ccggcttcca  | ttcagggtcga | gggtggcccg  | ctccatgcac  | cgcgacgcaa  | cgcggggagg  | 3960 |
| cagacaaggt  | atagggcggc  | gcctacaatc  | catgccaaac  | cggtccatgt  | gctcgccgag  | 4020 |
| goggcataaa  | tgcgctgac   | gatcagcggt  | ccagtgtatc  | aagttaggct  | ggtaagagcc  | 4080 |
| gcgagcgatc  | cttgaagctg  | tccctgatgg  | tcgtcatcta  | cctgcctgga  | cagcatggcc  | 4140 |
| tgcaacgcgg  | gcatcccgat  | gccgcccggaa | gcgagaagaa  | tcataatggg  | gaaggccatc  | 4200 |
| cagcctcgcg  | tcgcaaacgc  | cagcaagacg  | tagcccagcg  | cgctcgccgc  | catgcccctg  | 4260 |
| ttcatccccg  | tggtccggtg  | ctcgcggttg  | ctggcggtgt  | ccccggaaga  | aatatatttg  | 4320 |
| catgtcttta  | gttctatgat  | gacacaaaac  | ccgcccagcg  | tcttgctcatt | ggcgaaattcg | 4380 |
| aacacgcaga  | tgcaagtcgg  | gcggcgcggt  | cccaggtcca  | cttcgcatat  | taaggtgacg  | 4440 |
| cgtgtggcct  | cgaacaccga  | gcgaccctgc  | agcgacccgc  | ttaacagcgt  | caacagcgtg  | 4500 |
| ccgcagatcc  | cgggcaatga  | gatatgaaaa  | agcctgaact  | caccgacgac  | tctgtcgaga  | 4560 |
| agtttctgat  | cgaaaagtgc  | gacagcgtct  | ccgacctgat  | gcagctctcg  | gagggcgag   | 4620 |
| aatctcgtgc  | tttcagcttc  | gatgtaggag  | ggcgtggata  | tgtcctgcgg  | gtaaatagct  | 4680 |
| gcgcccagtg  | ttctacaaa   | gatcggtatg  | tttatcgga   | ctttgcatcg  | gcccgctcc   | 4740 |
| cgattccgga  | agtgttgac   | attggggaat  | tcagcgagag  | cctgacctat  | tgcatctccc  | 4800 |
| gccgtgcaca  | gggtgtcacg  | ttgcaagacc  | tgctgaaac   | cgaactgccc  | gctgttctgc  | 4860 |
| agccggtcgc  | ggaggccatg  | gatgcgatcg  | ctgcggccga  | tcttagccag  | acgagcggt   | 4920 |
| tcggcccat   | cggaccgcaa  | ggaatcggtc  | aatacactac  | atggcgatg   | ttcatatg    | 4980 |
| cgatttctga  | tcccctatgt  | tatcactggc  | aaactgtgat  | ggacgacacc  | gtcagtgctg  | 5040 |
| ccgtcgcgca  | ggctctcgat  | gagctgatgc  | tttgggcccga | ggactgcccc  | gaagtccggc  | 5100 |
| acctcggtga  | cgcggtattc  | ggctccaaca  | atgtcctgac  | ggacaatggc  | cgcataacag  | 5160 |
| cggctcattga | ctggagcgag  | gcgatgttcg  | gggatccca   | atacgaggtc  | gccaacatct  | 5220 |
| tcttctggag  | gccgtgggtg  | gcttgtagtg  | agcagcagac  | gcgctacttc  | gagcggaggc  | 5280 |
| atccggagct  | tgcaagatcg  | ccgcgggtcc  | gggcgtatat  | gctccgcatt  | ggtcttgacc  | 5340 |
| aactctatca  | gagcttggtt  | gacggcaatt  | tcgatgatgc  | agcttgggcg  | cagggctgat  | 5400 |
| gcgacgcaat  | cgtccgatcc  | ggagccggaa  | ctgtcgggcg  | tacacaaaac  | gcccgcgaa   | 5460 |
| gcggcgccgt  | ctggaccgat  | ggctgtgtag  | aagtactcgc  | cgatagtggg  | aaccgacgcc  | 5520 |
| ccagcactcg  | tcggagggca  | aaggaatagg  | ggagatgggg  | gaggctaact  | gaaacacgga  | 5580 |
| aggagacaat  | accggaagga  | acccgcgcta  | tgacggcaat  | aaaaagacag  | aataaaaacg  | 5640 |
| acgggtgttg  | ggtcggttgt  | tcataaacgc  | ggggttcggt  | cccagggtcg  | gcactctgtc  | 5700 |
| gataccccac  | cgagacccca  | ttggggccaa  | tacgcccgcg  | tttcttcctt  | ttccccaccc  | 5760 |
| caccccccaa  | gttcgggtga  | aggcccaggg  | ctcgagccca  | acgtcggggc  | ggcaggccct  | 5820 |
| gccatagcca  | ctggccccgt  | gggttaggga  | cggggtcccc  | catggggaa   | ggtttatggt  | 5880 |
| tcgtgggggt  | tattattttg  | ggcggttgcg  | ggggtctggt  | ccacgactgg  | actgagcaga  | 5940 |
| cagacccatg  | gtttttggat  | ggcctgggca  | tggaaccgat  | gtactggcgc  | gacacgaaca  | 6000 |
| ccgggctctc  | gtggctgcca  | aacacccccg  | acccccaaaa  | accaccgcgc  | ggatttctgg  | 6060 |
| cgccagtg    | cgtcgaccgg  | tcatggctgc  | gccccgacac  | ccgccaacac  | ccgctgacgc  | 6120 |
| gccctgacgg  | gcttgtctgc  | tcgccgcatc  | cgcttacaga  | caagctgtga  | ccgtctccgg  | 6180 |
| gagctgcatg  | tgtcagaggt  | tttcaccgtc  | atcacgaaa   | cgcgcgaggc  | agccggatca  | 6240 |
| taatcagcca  | taccacattt  | gtagaggttt  | tacttgcttt  | aaaaaacctc  | cccacctccc  | 6300 |
| cctgaacctg  | aaacataaaa  | tgaatgcaat  | tgttgttgtt  | aacttgttta  | ttgcagctta  | 6360 |
| taatggttac  | aaataaagca  | atagcatcac  | aaatttcaca  | aataaagcat  | ttttttcact  | 6420 |
| gcattctagt  | tgtgggttgt  | ccaaactcat  | caatgtatct  | tatcatgtct  | ggatccacta  | 6480 |
| gttctagagc  | ggccgcccac  | gcggtggagc  | tccagctttt  | gttcccttta  | gtgagggtta  | 6540 |
| atttcgagct  | tggtcgtaac  | atggctatag  | ctgtttcctg  | tgtgaaattg  | ttatccgctc  | 6600 |
| acaattccac  | acaacatacg  | agccggaagc  | ataaagtgtg  | aagcctgggg  | tgccaatga   | 6660 |
| gtgagtaaac  | tcacattaat  | tgcggtgcgc  | tcactgcccg  | ctttccagtc  | gggaaacctg  | 6720 |
| tcgtgcagc   | tgcattaatg  | aatcgcccaa  | cgcgcgggga  | gaggcggttt  | gcgtattggg  | 6780 |
| cgtctctccg  | cttccctcgt  | cactgactcg  | ctgcgctcgg  | tcgttcgggt  | gcggcgagcg  | 6840 |
| gtatcagctc  | actcaaaggc  | ggtaatacgg  | ttatccacag  | aatcagggga  | taacgcagga  | 6900 |
| aagaacatgt  | gagcaaaagg  | ccagcaaaa   | gccaggaacc  | gtaaaaaagg  | cgcggtgctg  | 6960 |
| gcgtttttcc  | ataggctccg  | ccccctgac   | gagcatcaca  | aaaatcgacg  | ctcaagtcag  | 7020 |
| aggtggcgaa  | acccgacagg  | actataaaga  | taccaggcgt  | ttccccctgg  | aagctccctc  | 7080 |
| gtgcgctctc  | ctggtccgac  | cctgccgctt  | accggatacc  | tgtccgcctt  | tctcccttcg  | 7140 |
| ggaagcgtgg  | cgctttctca  | tagctcacgc  | tgtaggatc   | tcagttcggt  | gtaggtcggt  | 7200 |
| cgctccaagc  | tgggctgtgt  | gcacgaaccc  | cccgttcage  | ccgaccgctg  | cgccttatcc  | 7260 |

```

ggtaactatc gtcttgagtc caaccggtg agacacgact tatcgccact ggcagcagcc 7320
actggtaaca ggattagcag agcgaggat gttagcggtg ctacagagtt cttgaagtgg 7380
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 7440
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 7500
gggtggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 7560
cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggatt 7620
ttggatcatg gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 7680
tttaaataca tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaatc 7740
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 7800
gtcgtgtaga taactacgat acgggaggggc ttaccatctg gccccagtgcc tgcaatgata 7860
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 7920
gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 7980
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacggtgt tgccattgct 8040
acaggcatcg tgggtgtcacg ctctcggttt ggtatggctt cattcagctc cggttcccaa 8100
cgatcaaggc gagttacatg atccccatg ttgtgcaaaa aagcgggttag ctccctcggt 8160
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 8220
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 8280
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 8340
atacgggata ataccgcgcc acatagcaga actttaaagg tgctcatcat tggaaaaagt 8400
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 8460
actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 8520
aaaacaggaa ggcaaaatgc cgcaaaaagg ggaataaggc cgacacggaa atgttgaata 8580
ctcatactct tctcttttca atattattga agcatttatc agggttattg tctcatgagc 8640
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 8700
cgaaaagtgc

```

<210> 5  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 5  
 atgggatcca agatgaagcg cgcaagaccg

30

<210> 6  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 6  
 cataacgcgg ccgcttcttt attcttgggc

30

<210> 7  
 <211> 7148  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: plasmid

<400> 7  
 gacggatcgg gagatctccc gatccccat ggtcgactct cagtacaatc tgctctgatg 60  
 ccgcatagtt aagccagtat ctgtccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120  
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcggttttgcg ctgcttcgcg atgtacgggc cagatatagc cgttgacatt 240  
gattattgac tagttatttaa tagtaatacaa ttacgggggc attagttcat agcccatata 300  
tcggagttccg cggttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360  
cccggccatt gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc 420  
attgacgtca atgggtggac tatttacggg aaactgccca cttggcagta catcaagtgt 480  
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600  
tcgctattac catgggtgatg cgggttttggc agtacatcaa tgggctgga tagcggtttg 660  
actcacgggg attttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720  
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggag 780  
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840  
ctgcttactg gcttatcgaa attaatagca ctactatag ggagacccaa gcttgggtacc 900  
gagctcggat ccaagatgaa gcgcgcaaga ccgtctgaag ataccttcaa ccccggtgat 960  
ccatagaca cggaaaccgg tccctcaact gtgecttttc ttactctctc ctttctatcc 1020  
cccaatgggt ttcaagagag tccccctggg ctactctctt tgcgcctatc cgaacctcta 1080  
gttacctcca atggcatgct tgcgctcaaa atgggcaacg gcctctctct ggacgagggc 1140  
ggcaacctta cctcccaaaa tgtaaccact gtgagccac ctctcaaaaa aaccaagtca 1200  
aacataaacc tggaaatata cgcggggcaac acagttacct cagaagccct aactgtggct 1260  
gcccgcgcac actccaaact tagcattgcc acccaaggac tgcaatcaca ggccccgcta 1320  
accgtgcacg actccaaact agcattgcc acccaaggac cctcacagt gtcagaagga 1380  
aagctagccc tgcaaacatc agggccccct accaccacg atagcagtag ccttactatc 1440  
actgcctcac cccctctaac tactgccact ggttagcttg gcattgactt gaaagagccc 1500  
atttatacac aaaatggaaa actaggacta aagtagcggt ccaggtgtga ctattaataa 1560  
gacctaaaca ctttgaccgt agcaactggg ggttagcttg gattcacaag gcaatatgca 1620  
caaactaaag ttactggagc cttgggtttt gatccacaag gcaatatgca acttaatgta 1680  
gcagccattac taaggattga ttctcaaaaac agacgcctta tacttgatgt tagttatccg 1740  
tttgatgctc aaaaccaact aaatctaaga ctaggacagg gccctctttt tataaactca 1800  
gcccacaact tggatattaa ctacaacaaa ggcctttact tgtttacagc ttcaaacaa 1860  
tccaaaaagc ttgaggttaa cctaagcact gccaaggggt tgatgtttga cgctacagcc 1920  
atagccatta atgcaggaga tgggcttgaa ttgggttcac ctaatgcacc aaacacaaat 1980  
cccctcaaaa caaaaattgg ccatggccta gaatttgatt caaacaaggc tatggttcct 2040  
aaactaggaa ctggccttag ttttgacagc acaggtgcca ttacagttag aaacaaaaat 2100  
aatgataagc taactttgtg gaccacacca gctccatctc ctaactgtag actaaatgca 2160  
gagaaagatg ctaaaactcac tttgggtctta acaaaatgtg gcagtcaaat acttgctaca 2220  
gtttcagttt tggctgttaa aggcagtttg gctccaatat ctggaacagt tcaaagtgtc 2280  
catcttatta taagatttga cgaataggga gttgactctc acaattcctt cctggacca 2340  
gaatattgga actttagaaa tggagatctt ggaagggca cagcctatac aaacgtgtt 2400  
ggatttatgc ctaacctatc agcttatcca aaatctcacg gtaaaactgc caaaagtac 2460  
attgtcagtc aagtttactt aaacggagac aaaactaaac ctgtaacact aaccattaca 2520  
ctaaacggta cacaggaac actccaagtg actcactctat catactctat gtcattttca 2580  
tgggactggg ctggccaca ctacattaat gaaatatttg ccacatctc ttacactttt 2640  
tcatacattg cccaagaata aagaagcggc cgctcgagca tgcactctaga gggccctatt 2700  
ctatagtgtc acctaaatgc tagagctcgc tgatcagcct cgactgtgcc ttctagtgtc 2760  
cagccatctg ttgtttgccc ctcccccggt ccttcttga ccttggaggg tggcactccc 2820  
actgtccttt cctaataaaa tgaggaaatt gcatcgcat gtctgagtag gtgtcattct 2880  
attctggggg gtgggttggg gcaggacagc aagggggagg attgggaaga caatagcagg 2940  
catgctgggg atgcggtggg ctctatggct tctgagcgcg cgttaagcg ctggggctct 3000  
aggggggtatc cccacgcgc cgttagcggc gcattaaagc cggcggggtg ggtggttacg 3060  
cgcagcgtga ccgctacact tgccagcgcc cttagcgccc ctcttttcgc tttcttccc 3120  
tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgagg catcccttta 3180  
gggttccgat ttagtgtttt acggcacctc ctgatagacg gtttttcgccc ggttgatggg 3240  
tcacgttagt ggccatcgcc ctgcatagac gttccaaact ggaacaacac ctttgacgtt 3300  
ttctttaata gtggactctt gttccaaact tccgcctatt tcaaccctat ctcggtctat 3360  
tcttttgatt tataagggat tttggggatt cgggctatt ggttaaaaaa tgagctgatt 3420  
taacaaaaat ttaacgcgaa ttaattctgt ggaatgtgtg tcagttaggg tgtggaaagt 3480  
ccccaggctc cccaggcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc 3540  
aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcattctaat 3600  
tagtcagcaa ctacagtccc gcccctaac cgcctcaac cgcctcaac tccgccagt 3660  
tccgcccatt ctccgcccc ctgctgacta atttttttta tttatgcaga ggccgaggcc 3720  
gcctctgcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt 3780  
tgcaaaaagc tccggggagc ttgtatatcc attttcggat ctgatcaaga gacaggatga 3840  
ggatcgtttc gcatgattga acaagatgga ttgcacgcag gttctcggc cgcttgggtg 3900

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gagaggctat  | tgggctatga  | ctgggacaaa  | cagacaatcg  | gctgctctga  | tgccgcccgtg | 3960 |
| ttccggctgt  | cagcgcaggg  | gcgcccgggt  | ctttttgtca  | agaccgacct  | gtccgggtgcc | 4020 |
| ctgaatgaac  | tgcaggacga  | ggcagcgcgg  | ctatcgtggc  | tggccacgac  | ggcggttcct  | 4080 |
| tgccgagctg  | tgctcgacgt  | tgctactgaa  | gcgggaagg   | actggctgct  | attgggcgaa  | 4140 |
| gtgccggggc  | aggatctcct  | gtcatctcac  | cttgctcctg  | ccgagaaagt  | atccatcatg  | 4200 |
| gctgatgcaa  | tgcggcgggt  | gcatacgctt  | gatccggcta  | cctgcccatt  | cgaccaccaa  | 4260 |
| gcgaaacatc  | gcatcgagcg  | agcacgtact  | cgatgggaag  | ccggctctgt  | cgatcaggat  | 4320 |
| gatctggacg  | aagagcatca  | ggggctcgcg  | ccagccgaac  | tgttcgccag  | gctcaaggcg  | 4380 |
| cgcctgcccc  | acggcgagga  | tctcgtcgtg  | acccatggcg  | atgcctgctt  | gccgaatatc  | 4440 |
| atgggtgaaa  | atggccgctt  | ttctggattc  | atcgactgtg  | gccggctggg  | tgtggcggag  | 4500 |
| cgctatcagg  | acatagcggt  | ggctaccogt  | gatattgctg  | aagagcttgg  | cggcgaatgg  | 4560 |
| gctgaccgct  | tcctcgtgct  | ttacgggtatc | gccgctcccg  | attcgagcg   | catcgccctc  | 4620 |
| tatcgccctc  | ttgacgaggt  | cttctgagcg  | ggactctggg  | gttcgaaatg  | accgaccaag  | 4680 |
| cgacgcccac  | cctgccatca  | cgagatttcg  | attccaccgc  | cgccctctat  | gaaaggttgg  | 4740 |
| gcttcggaat  | cgttttccgg  | gacgcggct   | ggatgatcct  | ccagcgcggg  | gatctcatgc  | 4800 |
| tggagttctt  | cgcccccccc  | aacttggtta  | ttgcagctta  | taatgggtac  | aaataaagca  | 4860 |
| atagcatcac  | aaatttcaca  | aataaagcat  | ttttttcact  | gcattctagt  | tgtggtttgt  | 4920 |
| ccaaactcat  | caatgtatct  | tatcatgtct  | gtataccgtc  | gacctctagc  | tagagcttgg  | 4980 |
| cgtaatcatg  | gtcatagctg  | tttctgtgt   | gaaattgtta  | tcgctcaca   | attccacaca  | 5040 |
| acatacgagc  | cggaagcata  | aagtgtaaag  | octggggtgc  | ctaagtgtg   | agctaactca  | 5100 |
| cattaattgc  | ggttgcgctca | ctgcccgtt   | tccagtcggg  | aaacctgtcg  | tgccagctgc  | 5160 |
| attaatgaat  | cggcccaacgc | gcggggagag  | gcggtttgcg  | tattggggcg  | tcctccgctt  | 5220 |
| cctcgctcac  | tgactcgctg  | cgctcggtcg  | ttcggctcgc  | gcgagcggta  | tcagctcact  | 5280 |
| caaaggcggg  | aatacgggtta | tcacacagaat | caggggataa  | cgaggaagaa  | aacatgtgag  | 5340 |
| caaaaaggcca | gcaaaaaggcc | aggaaccgta  | aaaaggccgc  | gttgctggcg  | ttttccata   | 5400 |
| ggctccgccc  | ccctgacgag  | catcacaaaa  | atcgacgctc  | aagtcagagg  | tggcgaaacc  | 5460 |
| cgacaggact  | ataaagatac  | caggcgtttc  | cccctggaag  | ctccctcgctg | cgctctcctg  | 5520 |
| ttccgaccct  | gccgcttacc  | ggatacctgt  | ccgcctttct  | cccttcggga  | agcggtggcg  | 5580 |
| tttctcaatg  | ctcacgctgt  | aggtatctca  | gttcgggtga  | ggctggtcgc  | tcgaagctgg  | 5640 |
| gctgtgtgca  | cgaaccccc   | gttcagcccg  | accgctgcgc  | cttatccggt  | aactatcgct  | 5700 |
| ttgagtccaa  | cccggtaaga  | cacgacttat  | cgccactggc  | agcagccact  | ggtaacagga  | 5760 |
| ttagcagagc  | gaggtatgta  | ggcggtgcta  | cagagtctct  | gaagtgggtg  | cctaactacg  | 5820 |
| gctacactag  | aaggacagta  | tttggtatct  | gcgctctgct  | gaagccagtt  | accttcggaa  | 5880 |
| aaagagttgg  | tagctcttga  | tcgggcaaac  | aaaccaccgc  | tggtagcggg  | ggtttttttg  | 5940 |
| tttgcaagca  | gcagattacg  | cgcagaaaaa  | aaggatctca  | agaagatcct  | ttgatctttt  | 6000 |
| ctacggggct  | tgacgctcag  | tggaaacgaa  | actcacgtta  | agggattttg  | gtcatgagat  | 6060 |
| tatcaaaaag  | gatcttcacc  | tagatccttt  | taaattaaaa  | atgaagtttt  | aatcaatct   | 6120 |
| aaagtatata  | tgagtaaaact | tggtctgaca  | gttaccaatg  | cttaatcagt  | gaggcaccta  | 6180 |
| tctcagcgat  | ctgtctatct  | cgttcatcca  | tagttgcctg  | actcccgcgc  | gtgtagataa  | 6240 |
| ctacgatacg  | ggaggggctta | ccatctggcc  | ccagtgtgct  | aatgataccg  | cgagaccac   | 6300 |
| gctcaccggc  | tccagattta  | tcagcaataa  | accagccagc  | cggaagggcc  | gagcgagaa   | 6360 |
| gtggctcctg  | aacttttatcc | gcctccatcc  | agtctattaa  | ttgttgccgg  | gaagctagag  | 6420 |
| taagtagttc  | gccagttaat  | agtttgcgca  | acgttggtgc  | catgtctaca  | ggcatcggtg  | 6480 |
| tgctacgctc  | gtcggttggt  | atggcttcat  | tcagctccgg  | ttcccaacga  | tcaaggcgag  | 6540 |
| ttacatgatc  | ccccatggtg  | tgcaaaaaag  | cggttagctc  | cttcgggtcct | ccgatcggtg  | 6600 |
| tcagaagtaa  | ggtggccgca  | gtgttatcac  | tcattggttat | ggcagcactg  | cataattctc  | 6660 |
| ttactgtcat  | gccatccgta  | agatgctttt  | ctgtgactgg  | tgagtactca  | accaagtcat  | 6720 |
| tctgagaata  | gtgtatgctg  | cgaccgagtt  | gctcttgccc  | ggcgtcaata  | cgggataata  | 6780 |
| ccgcgccaca  | tagcagaact  | ttaaaagtgc  | tcattattgg  | aaaacgttct  | tcggggcgaa  | 6840 |
| aactctcaag  | gatcttaccg  | ctgttgagat  | ccagttcgat  | gtaaccact   | cgtgcacca   | 6900 |
| actgatcttc  | agcatctttt  | actttcacca  | gcgtttctgg  | gtgagcaaaa  | acaggaaggc  | 6960 |
| aaaatgccgc  | aaaaaaggga  | ataaggcgga  | cacggaaatg  | ttgaatactc  | atactcttcc  | 7020 |
| tttttcaata  | ttattgaagc  | atttatcagg  | gttattgtct  | catgagcgga  | tacatatttg  | 7080 |
| aatgtattta  | gaaaaataaa  | caaatagggg  | ttccgcgcac  | atttccccga  | aaagtgccac  | 7140 |
| ctgacgtc    |             |             |             |             |             | 7148 |

&lt;210&gt; 8

&lt;211&gt; 7469

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid

&lt;400&gt; 8

```

gacggatcgg gagatctccc gatcccttat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gtagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttaggggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaataaa ttacgggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggccccgcc tggctgaccg cccaacgacc 360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgcccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg cttatcgaa cttaatacga ctactatag ggagacccaa gcttggtagc 900
gagctcggat ctgaattcga gctcgctgtt gggctcgcgg ttgaggacaa actcttcgcg 960
gtctttccag tactcttggg tccgaaaacc gtcggcctcc gaacggtagt ccgccaccga 1020
gggacctgag cgagtccgca tcgaccggat cggaaaacct ctcgagaaag gcgtctaacc 1080
agtcacagtc gcaaggtagg ctgagcaccg tggcggggcg cagcgggtgg cggctcgggg 1140
tggtttctggc ggaggtgctg ctgatgatgt aattaaagta ggcgtcttg agacggcgga 1200
tggtcgaggt gaggtgtggc aggtctgaga tccaagatga agcgcgcaag accgtctgaa 1260
gataccttca acccggtgta tccatatgac acggaaaacc gtccctcaac tgtgcctttt 1320
cttactcttc cctttgtatc cccaatggg tttcaagaga gtccccctgg ggtactctct 1380
ttgcgcctat ccgaacctct agttacctcc aatggcatgc ttgcgctcaa aatgggcaac 1440
ggcctctctc tggacgaggg cggcaacctt acctcccaaa atgtaaccac tgtgagccca 1500
cctctcaaaa aaaccaagtc aaacataaac ctggaaatat ctgcaccctt cacagttacc 1560
tcagaagccc taactgtggc tgccgcccga cctctaattg tcgcgggcaa cacactcacc 1620
atgcaatcac aggccccgct aaccgtgcac gactccaaac ttagcattgc caccacaagg 1680
cccctcacag tgtcagaagg aaagctagcc ctgcaaacat caggccccct caccaccacc 1740
gatagcagta cccttactat cactgcctca cccctctaa ctactgccac tggtagcttg 1800
ggcattgact tgaagagacc catttatata caaaatggaa aactaggact aaagtacggg 1860
gctcctttgc atgtaacaga cgacctaaac actttgaccg tagcaactgg tccaggtgtg 1920
actattaata atacttctt gcaaaactaa gttactggag ccttgggttt tgattcacia 1980
ggcaatatgc aacttaatgt agcaggagga ctaaggattg attctcaaaa cagacgcctt 2040
atacttgatg ttagttatcc gtttgatgct caaaaaccaac taaatctaag actaggacag 2100
ggcctctttt ttataaactc agcccacaac actacaaaca actacaacaa aggcctttac 2160
ttgtttacag cttcaaacaa ttccaaaagg cttgaggtta acctaagcac tgccaagggg 2220
ttgatgtttg acgtacagc catagccatt aatgcaggag atgggcttga atttggttca 2280
cctaattgcac caaacacaaa tcccctcaaa actggcctta gttttgacag cacaggtgcc 2340
tcaaacaggg ctatgggttc taaactagga actggcctta gttttgacag cacaggtgcc 2400
attacagtag gaaacaaaaa taatgataag ctaactttgt ggaccacacc agctccatct 2460
cctaactgta gactaaatgc agagaaagat gctaaactca ctttggctct aacaaaatgt 2520
ggcagtcгаа tacttgctac agtttcagtt ttggctgtta aaggcagttt ggctccaata 2580
tctggaacag ttcaaagtgc tcatcttatt ataagatttg acgaaaatgg agtgctacta 2640
aacaattcct tcctggaccc agaattattg aactttagaa atggagatct tactgaaggc 2700
acagcctata caaacgctgt tggatttatg cctaacctat cagcttatcc aaaattctac 2760
ggtaaaactg ccaaaaagtaa cattgtcagt caagtttact taaacggaga caaaactaaa 2820
cctgtaacac taaccattac actaaacggt acacaggaaa caggagacac aactccaagt 2880
gcatactcta tgtcattttc atgggactgg tctggccaca actacattaa tgaaatattt 2940
gccacatcct cttacacttt ttcatacatt gcccagaagt aaagaagcgg ccgctcgagc 3000
atgcactctag agggccctat tctatagtgt cacctaaatg ctagagctcg ctgatcagcc 3060
tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctcccccggt gccttccttg 3120
accctggaag gtgccactcc cactgtcctt tccctaataa atgaggaaat tgcctcgcat 3180
tgtctgagta ggtgtcattc tattctgggg ggtgggggtg ggcaggagac caagggggag 3240
gattgggaag acaatagcag gcatgctggg gatgcggttg gctctatggc ttctgaggcg 3300
gaaagaacca gctggggctc tagggggtat cccacgcgc cctgtagcgg cgcatgaagc 3360
gcgggcgggtg tgggtggttac gcgcagcgtg accgtacac ttgccagcgc ctagcgccc 3420
gctcctttcg ctttcttccc ttccctttctc gccacgttcg ccggctttcc ccgtcaagct 3480

```

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ctaaatcggy  | gcatcccttt  | agggttccga  | tttagtgctt  | tacggcacct  | cgacccccaa  | 3540 |
| aaacttgatt  | agggtgatgg  | ttcacgtagt  | gggccatcgc  | cctgatagac  | ggtttttcgc  | 3600 |
| cctttgacgt  | tggagtccac  | gttctttaat  | agtggactct  | tggtccaaac  | tggaaacaaca | 3660 |
| ctcaacccta  | tctcggctta  | ttcttttgat  | ttataaggga  | ttttggggat  | ttcggcctat  | 3720 |
| tggttaaaaa  | atgagctgat  | tttaacgcga  | tttaacgcga  | attaattctg  | tggaaatgtg  | 3780 |
| gtcagttagg  | gtgtggaaag  | ttcccaggct  | ccccaggcag  | gcagaagtat  | gcaaagcatg  | 3840 |
| catctcaatt  | agt cagcaac | cagggtgtgga | aagtccccag  | gctccccagc  | aggcagaagt  | 3900 |
| atgcaaaaga  | tgcattctca  | ttagtcagca  | accatagtcc  | cgccccctaac | tccgccccatc | 3960 |
| cgcggccctaa | ctccgcccag  | ttccgcccag  | tctccgcccc  | atggctgact  | aattttttttt | 4020 |
| atztatgcag  | aggccgaggc  | cgcctctgcc  | tctgagctat  | tccagaagta  | gtgaggaggc  | 4080 |
| ttttttggag  | gcctaggctt  | ttgcaaaaag  | ctcccggggag | cttgatatatc | catttttcgga | 4140 |
| tctgatcaag  | agacaggatg  | aggatcggtt  | cgcatgattg  | aacaagatgg  | attgcacgca  | 4200 |
| gggtctccgg  | ccgcttgggt  | ggagaggcta  | ttcggctatg  | actgggcaca  | acagacaatc  | 4260 |
| ggctgctctg  | atgcgcgctg  | gttccggctg  | tcagcgcagg  | ggcgccccgt  | tctttttgtc  | 4320 |
| aagaccgacc  | tgtccgggtgc | cctgaatgaa  | ctgcaggacg  | aggcagcgcg  | gctatcgtgg  | 4380 |
| ctggccacga  | cggggttcc   | ttgcgcagct  | gtgctcgacg  | ttgtcactga  | agcgggaagg  | 4440 |
| gactggctgc  | tattgggcga  | agtgcggggg  | caggatctcc  | tgtcatctca  | ccttgctcct  | 4500 |
| gccgagaaag  | tatccatcat  | ggctgatgca  | atgcggcgcc  | tgcatacgtt  | tgatccggct  | 4560 |
| acctgcccac  | tcgaccacca  | agcgaaacat  | cgcctcgagc  | gagcacgtac  | tcggatggaa  | 4620 |
| gccggctctt  | togatcagga  | tgatctggac  | gaagagcatc  | aggggctcgc  | gccagccgaa  | 4680 |
| ctgttcgcca  | ggctcaaggc  | gcgcagtcgc  | gacggcgagg  | atctcgctcg  | gacccatggc  | 4740 |
| gatgcctgct  | tgccgaatat  | catgggtgaa  | aatggccgct  | tttctggatt  | catcgactgt  | 4800 |
| ggcggctctg  | gtgtggcgga  | ccgctatcag  | gacatagcgt  | tggctaccog  | tgatattgct  | 4860 |
| gaagagcttg  | gcggcgcaatg | ggctgaccgc  | ttcctcgtgc  | tttacggtat  | cgccgctccc  | 4920 |
| gattcgcagc  | gcacgcgctt  | ctatcgctt   | cttgacgagt  | tcttctgagc  | gggactctgg  | 4980 |
| ggttcgaaat  | gaccgaccaa  | gcgacgcccc  | acctgcccac  | acgagatttc  | gattccaccg  | 5040 |
| ccgcctctca  | tgaagggttg  | ggcttcggaa  | tcgttttcog  | ggacgcgggc  | tggatgatcc  | 5100 |
| tccagcgcg   | ggatctcatg  | ctggagttct  | tcgcccaccc  | caacttggtt  | attgcagctt  | 5160 |
| ataatggtta  | caaataaagc  | aatagcatca  | caaatttcac  | aaataaagca  | tttttttcac  | 5220 |
| tgcattctag  | ttgtgggttg  | tccaaaactca | ttcatgtatc  | ttatcatgtc  | tgtataccgt  | 5280 |
| cgacctctag  | ctagagcttg  | gcgtaatcat  | ggctcatagct | gtttcctgtg  | tgaattgttt  | 5340 |
| atccgctcac  | aattccacac  | aacatacag   | ccggaagcat  | aaagtgtaaa  | gcctgggggtg | 5400 |
| cctaagtagt  | gagctaactc  | acattaattg  | cgttgcgctc  | actgcccgc   | ttccagtcgg  | 5460 |
| gaaacctgtc  | gtgccagctg  | cattaatgaa  | tcggcccaacg | cgcggggaga  | ggcggtttgc  | 5520 |
| gtattgggcg  | ctcttcgct   | tcctcgctca  | ctgactcgct  | gcgctcggtc  | gttcggctgc  | 5580 |
| ggcgagcggt  | atcagctcac  | tcaaaaggcg  | taatacgggt  | atccacagaa  | tcaggggata  | 5640 |
| acgcaggaaa  | gaacatgtga  | gcaaaaggcc  | agcaaaaggc  | caggaaccgt  | aaaaaggccg  | 5700 |
| cggttctggc  | gtttttccat  | aggctccgcc  | cccccgacga  | gcatacaaaa  | aatcgacgct  | 5760 |
| caagtccagag | gtggcgaaac  | ccgacaggac  | tataaagata  | ccaggcggtt  | ccccctggaa  | 5820 |
| gctccctcgt  | gcgctctcct  | gttccgaccc  | tgccgcttac  | cggataacctg | tccgcctttc  | 5880 |
| tcccttcggg  | aagcgtggcg  | ctttctcaat  | cttcacgctg  | taggtatctc  | agttcggtgt  | 5940 |
| aggctcgttcg | ctccaagctg  | ggctgtgtgc  | acgaaccccc  | cgttcagccc  | gaccgctgcg  | 6000 |
| ccttatccgg  | taactatcgt  | cttgagtcca  | acccggttaag | acacgactta  | tcgccactgg  | 6060 |
| cagcagccac  | tggttaacagg | attagcagag  | cgaggtatgt  | aggcgggtgct | acagagttct  | 6120 |
| tgaagtgggtg | gcctaactac  | ggctacacta  | gaaggacagt  | atttgggtatc | tgcgctctgc  | 6180 |
| tgaagccagt  | taccttcgga  | aaaagagttg  | gtagctcttg  | atccggcaaaa | caaaccaccg  | 6240 |
| ctggtagcgg  | tgggtttttt  | gtttgcaagc  | agcagattac  | gcgcagaaaa  | aaaggatctc  | 6300 |
| aagaagatcc  | tttgatcctt  | tctacggggt  | ctgacgctca  | gtggaacgaa  | aactcacggt  | 6360 |
| aagggttttt  | ggctatgaga  | ttatcaaaaa  | ggatcttcac  | ctagatcctt  | ttaaattaaa  | 6420 |
| aatgaagttt  | taaatcaatc  | taaagtatat  | atgagtaaac  | ttggctctgac | agttaccaat  | 6480 |
| gcttaatcag  | tgaggcacct  | atctcagcga  | tctgtctatt  | tcgttcaccc  | atagttgcct  | 6540 |
| gactccccgt  | cgtgtagata  | actacgatac  | gggaggggctt | accatctggc  | cccagtgctg  | 6600 |
| caatgatacc  | gcgagaccca  | cgctcaccgg  | ctccagattt  | atcagcaata  | aaccagccag  | 6660 |
| ccggaagggc  | cgagcgcaga  | agtggctcctg | caactttatc  | cgcctccatc  | cagtctatta  | 6720 |
| attggtgccc  | ggaagctaga  | gtaagtagtt  | cgccagttaa  | tagtttgccg  | aacgttggtg  | 6780 |
| ccattgctac  | aggcatcggt  | gtgtcacgct  | cgtcggttgg  | tatggcttca  | ttcagctccg  | 6840 |
| gttcccaacg  | atcaaggcga  | gttacatgat  | ccccatggt   | gtgcaaaaaa  | gcgggttagct | 6900 |
| ccttcgggtc  | tccgatcggt  | gtcagaagta  | agtggccgcg  | agtgttatca  | ctcatgggta  | 6960 |
| tggcagcaat  | gcataattct  | cttactgtca  | tgccatccgt  | aagatgcttt  | tctgtgactg  | 7020 |
| gtgagtactc  | aaccaagtca  | ttctgagaat  | agtgtatgcg  | gcgaccgagt  | tgtctctgcc  | 7080 |
| cggcgctcaat | acgggataat  | accgcgccac  | atagcagaac  | tttaaaagtg  | ctcatcattg  | 7140 |
| gaaaacgttc  | ttcggggcga  | aaactctcaa  | ggatcttacc  | gctgttgaga  | tccagttcga  | 7200 |

tgtaaccac tctgtcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 7260  
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 7320  
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 7380  
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 7440  
catttccccg aaaagtgcc cctgacgtc 7469

<210> 9  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 9  
tgcttaagcg gccgcgaagg agaagtcc 28

<210> 10  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 10  
ccgagctagc gactgaaaat gag 23

<210> 11  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 11  
cctctcgaga gacagcaaga cac 23

<210> 12  
<211> 11152  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: plasmid

<400> 12  
aagcttgggc agaaatgggt gaactcccga gagtgtccta cacctagggg agaagcagcc 60  
aaggggttgt ttcccaccaa ggacgaccgg tctgcgccaca aacggatgag cccatcagac 120  
aaagacatat tcattctctg ctgcaaaactt ggcatagctc tgctttgcct ggggctattg 180  
ggggaagtgt cggttcgtgc tcgcagggct ctacaccttg actcttttaa tagctcttct 240  
gtgcaagatt acaatctaaa caattcggag aactcgacct tcctcctgag gcaaggacca 300  
cagccaactt cctcttacaa gccgcacatga ttttgtcctt cagaaataga aataagaatg 360  
cttgctaaaa atttatattt taccaataag accaatccaa taggtagatt attagttact 420  
atgttaagaa atgaatcatt atcttttagt actattttta ctcaaatca gaagttagaa 480  
atgggaatag aaaatagaaa gagacgctca acctcaattg aagaacagggt gcaaggacta 540  
ttgaccacag gcctagaagt aaaaaagggg aaaaagagtg tttttgtcaa aataggagac 600  
aggtgggtgc aaccagggac ttatagggga cttacatct acagaccaac agatgcccc 660  
ttaccatata caggaagata tgacttaaat tgggataggt ggggtacagt caatggctat 720

aaagtgttat atagatccct cccttttcgt gaaagactcg ccagagctag acctccttgg 780  
tgatatgtgt ctcaagaaga aaaagacgac atgaaacaac aggtacatga ttatatattat 840  
ctaggaacag gaatgcactt ttggggaaag attttccata ccaaggaggg gacagtggct 900  
ggactaatag aacattattc tgcaaaaact catggcatga gttattatga atagccttta 960  
ttggcccaac cttgcggttc ccagggttta agtaagtttt tgggtacaaa ctgttcttaa 1020  
aacgacaagt tgagacaagt ggtttcctga cttggtttgg tatcaaaggt tctgatctga 1080  
gctctgagtg ttctattttc ctatgttctt ttggaattta tccaaatctt atgtaaatgc 1140  
ttatgtaaac caagatataa aagagtgcctg attttttgag taaacttgca acagtcctaa 1200  
cattcacctc ttgtgtgttt gtgtctgttc gccatcccg ctcgcctcgt cacttatcct 1260  
tcactttcca gaggtcccc ccgcagacct cggcgacct caggctcgcc gactcgccga 1320  
gctggcgccc gaacagggac cctcggataa gtgaccttg tctctatttc tactatttgg 1380  
tgtttgtctt gtattgtctc tttcttgtct ggctatcatc acaagagcgg aacggactca 1440  
ccatagggag caagctagcg actgaaaatg agacatatta tctgccacgg aggtgttatt 1500  
accgaagaaa tggccgccag tcttttggac cagctgatcg aagaggtact ggctgataat 1560  
cttccacctc cttagccattt tgaaccacct acccttcacg aactgtatga ttttagacgtg 1620  
acggcccccg aagatcccaa cgaggaggcg gtttcgcaga tttttcccg cttctgtaatg 1680  
ttggcggtgc aggaagggat tgacttactc acttttccgc cggcgcccg ttctccggag 1740  
ccgcctcacc tttccggca gcccgagcag ccggagcaga gagccttggg tccggtttct 1800  
atgccaaacc ttgtaccgga ggtgatcgat cttacctgcc acgaggctgg ctttccaccc 1860  
agtgcgacg aggatgaaga gggtagggag tttgtgttag attatgtgga gcaccccggg 1920  
cacggttgca ggtcttgtca ttatcacgg aggaatacgg gggaccaga tattatgtgt 1980  
tcgctttgct atatgaggac ctgtggcatg tttgtctaca gtaagtgaat attatgggca 2040  
gtgggtgata gagtgggtggg tttggtgtgg taattttttt ttttaatttt acagttttgt 2100  
ggtttaaaga attttgtatt gtgatttttt taaaagggtc tgtgtctgaa cctgagcctg 2160  
agccccagcc agaaccggag cctgcaagac cctgcagcag cctaccccg tcctaaaatg gcgcctgcta 2220  
tcctgagacg cccgacatca cctgtgtcta gagaatgcaa tagtagtacg gatagctgtg 2280  
actccggtcc ttctaacaca cctcctgaga tacacccggt ggtcccgctg tgccccatta 2340  
aaccagtgc cgtgagagt gggtggcgct gccaggctgt ggaatgtatc gaggacttgc 2400  
ttaacgagcg tgggcaacct ttggacttga gctgtaaacg cccagggcca taaggtgtaa 2460  
acctgtgatt gcgtgtgtgg ttaacgcctt tgtttgtcga atgagttgat gtaagtttaa 2520  
taaagggtga gataatgttt aacttgcctg gcgtgttaaa tggggcgggg cttaaagggt 2580  
atataatgcg cctgggcta atcttgggtt catctgacct catggaggct tgggagtgtt 2640  
tggaagattt ttctgtgtg cgtaacctgc tggaacagag ctctaacagt acctcttgg 2700  
tttggagggt tctgtggggc tcatcccagg caaagttagt ctgcagaatt aaggaggatt 2760  
acaagtggga atttgaagag cttttgaaat cctgtgtgta gctgtttgat tctttgaatc 2820  
tgggtcacca ggcgttttcc caagagaagg tcatcaagac tttggatttt tccacacggg 2880  
ggcgcgctgc ggctgtgtt gcttttttga gttttataaa ggataaatgg agcgaagaaa 2940  
cccactgtag cgggggggtac ctgtgtgatt ttctggccat gcatctgtgg agagcggttg 3000  
tgagacacaa gaactgcctg ctactgttgt cttcogtccg cccggcgata ataccgacgg 3060  
aggagcagca gcagcagcag gaggaagcca gcggcgcgcg gcaggagcag agcccattgga 3120  
acccgagagc cggcctggac cctcgggaat gaattgttga cagggtggctg aactgtatcc 3180  
agaactgaga cgcattttga caattacaga ggatgggcag gggctaaagg gggtaagag 3240  
ggagcggggg gcttgtgagg ctacagagga ggctaggaat ctagctttta gcttaatgac 3300  
cagacaccgt cctgagtgtt ttacttttca acagatcaag gataattgag ctaatgagct 3360  
tgatctgtct gcgcagaagt attccataga gcagctgacc acttactggc tgcagccagg 3420  
ggatgatttt gaggaggcta ttaggtata tgcaaagggt gcacttaggc cagattgcaa 3480  
gtacaagatc agcaaaactt taaatcacg gaattgttgc tacatttctg ggaacggggc 3540  
cgaggtggag atagatacgg aggatagggt ggcctttaga tgtagcatga taaatatgtg 3600  
gccgggggtg cttggcatgg acgggggtgt tattatgaat gtaaggttta ctggcccca 3660  
ttttagcggg acggttttcc tggccaatac caaccttatc ctacacgggt taagcttcta 3720  
tgggtttaac aatacctgtg tgggaagcctg gaccgatgta agggttcggg gctgtgcctt 3780  
ttactgtctg tgggaagggg ttgtgtgtcg ccccaaaagc agggcttcaa ttaagaaatg 3840  
cctctttgaa aggtgtacct tgggtatcct gtctgagggt aactccaggg tgcgccacaa 3900  
tgtggcctcc gcttcatgct gcttcaaaagc gtggctgtga ttaagcataa 3960  
catggtatgt ggcaactgcg aggacagggc ctctcagatg ctgacctgct cggacggcaa 4020  
ctgtcacctg ctgaagacca ttcacgtagc cagccactct cgcaaggcct ggccagtggt 4080  
tgagcataac atactgccc gctgttctt gcatttgggt aacaggaggg ggggtgttct 4140  
accttaccaa tgcaatttga gtcacactaa gatattgctt gagcccgaga gcatgtccaa 4200  
gggtgaacctg aacggggtgt ttgacatgac catgaagatc tggaaggtgc tgaggtacga 4260  
tgagacccgc accaggtgca gaccctgcga gtgtggcggt aaacatatta ggaaccagcc 4320  
tgtgatgctg attgtgaccg aggagctgag gcccgatcac ttgggtgctg cctgcaccg 4380  
cgctgagttt ggctctagcg atgaagatc agattgaggt actgaaatgt gtggggctgg 4440



```

cttaaggggtg ggaaagaata tataaggtgg ggggtcttatg tagtttttga tctgtttttgc 4500
agcagccgccc gccgccatga gcaccaactc gtttgatgga agcattgtga gctcatatatt 4560
gacaacgcgc atgcccccat gggccgggggt gcgtcagaat gtgatgggct ccagcattga 4620
tggtcgcccc gtcctgcccc caaactctac taccttgacc tacgagaccg tgtctgggaa 4680
gccgtttggag atcgcagcct ccgccgcgcg ttcagccgct gcagccaccg cccgcgggat 4740
tgtgactgac tttgctttcc tgagcccgct tgcaagcagt gcagcttccc gttcatccgc 4800
ccgcgatgac aagttgacgg ctctttttggc acaattggat tctttgaccc gggaaacttaa 4860
tgtcgtttct cagcagctgt tggatctgcg ccagcagggt tctgcccctga aggttccctc 4920
ccctcccaat gcggtttaaa acataaataa aaaaccagac tctgtttgga tttggatcaa 4980
gcaagtgtct tgctgtctct cgagggatct acagagattt aaagctctaa ggtaaatata aaatttttaa 5040
cataattgga caaactacct ctgattctaa ttgtttgtgt attttagatt ccaacctatg 5100
gtgtataatg tgtaaacta atgggagcag tggtggaatg cctttaatga ggaaaacctg ttttgctcag 5160
gaactgatga atctagtgat gatgaggata ctgctgactc tcaacattct actctccaa 5220
aagaaatgcc aaaggtagaa gaccccaagg actttccctc agaattgcta agttttttga 5280
aaaagaagag gtttagtaat agaactcttg cttgctttgc tatttacacc acaaaggaaa 5340
gtcatgctgt gctatacaag aaaattatgg aaaaatattc tgtaaccttt ataagtaggc 5400
aagctgcact taatcataac atactgtttt ttcttactcc acacaggcat agagtgtctg 5460
ataacagtta ctatgctcaa aaattgtgta ccttttagct ttttaattgt aaaggggtta 5520
ctattaataa tttgatgtat agtgccctga ctagagatca taatcagcca taccacattt 5580
ataaggaata tttgatgtat agtgccctga ctacacctcc cctgaacct gaaacataaa 5640
gtagaggttt tacttgcttt aaaaaacctc ccacacctcc attgcagctt ataatgggta caaataaagc 5700
atgaatgcaa ttgttggtgt taacttggtt tttttttcac tgcattctag ttgtgggttg 5760
aatagcatca caaatttcac aaataaagca ttgatccggt tggtggaatgt gtgtcagtta 5820
tccaaactca tcaatgtatc ctccccagca ggacagaagta tgcaaaagcat gcatctcaat 5880
gggtgtggaa agtccccagg ccaggtgtgg aaagtcccca ggctccccag caggcagaag 5940
tagtcagcaa ccaggtgtgg aacctatgtc ccgccccctaa ctccgccccat cccgccccct 6000
atgcatctca attagtcagc gttccgcccc ctctgagcta ttccagaagt agtgaggagg 6060
actccgcccc cgcctcggc tttgcaaaaa gcttgagcac aagacagggt tgcgagatat gtttgagaat 6120
gaggccgagg cccgctcggc tttgcaaaaa gcttgagcac aagacagggt tgcgagatat gtttgagaat 6180
ggcctaggct cccgctcggc tttgcaaaaa gcttgagcac aagacagggt tgcgagatat gtttgagaat 6240
accactttat cccgctcggc tttgcaaaaa gcttgagcac aagacagggt tgcgagatat gtttgagaat 6300
actgggtttt agtgcccgag gtagataaac aggtcgggac atctctataa cctattttcc cctcgaacac 6360
tttttaagcc gtagataaac aggtcgggac atctctataa cctattttcc cctcgaacac 6420
ctgggacatg ttgcagatcc atgcacgtaa actcgcgaagc cgactgatgc cttctgaaca 6480
atggaaaggc attattgccc taagccgtgg tgcgtatgga cgggtgagc agctacgac acgacaacca 6540
tgaactgggt attcgatcag tgcataccgt aggtcagatg ggaaggttca tcgttattga 6600
gcgcgagctt aaagtgtgta aacgcgcaga aggcgatggc gaaatgtatc caaaagcgca 6660
tgacctgggt gataccgggt gtactgcggt tgcgattcgt gttgatgact atgttgttga 6720
ctttgtcacc atcttcgcaa aaccggctgg tcgtccgctg gttgatgact atgttgttga 6780
tatcccgcaa gatacctgga ttgaacagcc gtgggatatg ggcgtcgtat tcgtcccgcc 6840
aatctccggt cgctaattct ttcaacgcct ttaaacatcc tgaaacctcg acgctagtcc 6900
caggcggggt acaatagttt aaaccccgct ttaaacatcc tgaaacctcg acgctagtcc 7020
acctgagcga aacctgttcc caaccgcctg tgcatcgcc ccttgatggg aaaacctcc 7080
gccgctttaa tcacggcgca caaccgcctg tgcatcgcc ccttgatggg aaaacctcc 7140
ctcactggga tcgcatgatt aaccgtctga tggatctctg gcgcggcatt gacccacgcg 7200
aaatcctcga cgtccaggca cgtatttgta tgagcgatgc cgaacgtacc gacgatgatt 7260
tatacgatac ggtgatggc taccgtggcg gcaactggat ttatgagtgg gccccggatc 7320
tttgtgaagg aaccttactt ctgtgggtgtg acataattgg gtgtataat gtgttaaact actgattcta 7380
taaagctcta aggtaaatat aaaattttta agtgatataa aatgggagca gtgggtggaat 7440
attgtttgtg tatttttagat tccaacctat gttttgctca gaagaaatgc catctagtga tgatgaggct 7500
gcctttaatg aggaaaaacct tactcctcca aaaaagaaga gaaaggtaga agaccccaag 7560
actgctgact ctcaacattc tactcctcca aaaaagaaga gaaaggtaga agaccccaag 7620
gactttcctt cagaattgct aagttttttg agtcactgtg tgctatacaa gaaaattatg 7680
gcttgctttg ctatttacac cacaaggaa aaagctgcac ataactcata catactgttt 7740
gaaaaatatt ctgtaacctt cacacaggca tagagtgtct gctattaata actatgctca aaaattgtgt 7800
tttcttactc ttttaatttg taaaggggtt aataagggaat atttgatgta tagtgccctg 7860
acctttagct ttttaatttg taaaggggtt aataagggaat atttgatgta tagtgccctg 7920
actagagatc ataactcagc ataccacatt tgtagagggt ttacttgctt taaaaaacct 7980
cccacacctc cccctgaacc tgaacataa aatgaatgca attgttgggt ttaacttggg 8040
tattgcagct tataatggtt acaataaag gttgtgggtt gtccaaactc atcaatgtat cttatcatgt 8100
atttttttca ctgcattcta gttgtgtcct cataaacctc aacctcctc acttgagagg 8160
ctggatcccc aggaagctcc tctgtgtcct cataaacctc aacctcctc acttgagagg 8160

```

acattccaat cataggctgc ccattccacc tctgtgtcct cctgttaatt aggtcactta 8220  
acaaaaagga aattgggtag ggggtttttca cagaccgctt tctaagggtta attttaaaat 8280  
atctgggaag tcccttccac tgctgtgttc cagaagtgtt ggtaaacagc ccacaaatgt 8340  
caacagcaga aacatacaag ctgtcagctt tgcacaaggg cccaacacc tgcctcatca 8400  
gaagcactgt ggttgctgtg ttagtaatgt gcaaaacagg aggcacattt tccccacctg 8460  
tgtaggttcc aaaatatcta gtgttttcat ttttacttgg atcaggaacc cagcactcca 8520  
ctggataagc attatcotta tccaaaacag ccttgtggtc agtgttcate tgctgactgt 8580  
caactgttagc attttttggg gttacagttt gagcaggata tttgggtcctg tagtttgcta 8640  
acacaccctg cagctccaaa ggttccccac caacagcaaa aaaatgaaaa tttgacctt 8700  
gaatgggttt tccagcacca ttttcatgag tttttgtgt ccctgaatgc aagttaaaca 8760  
tagcagttac cccaataacc tcagttttta cagtaacagc tccccacatc aaaatatttc 8820  
cacagggtta gtcctcattt aaattaggca aaggaattct tgaagacgaa agggcctcgt 8880  
gatacgctta tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcaggtgg 8940  
cacttttctg ggaatgtgc gcggaacccc tatgtgttta ttttctaaa tacattcaaa 9000  
tatgtatcgc ctcagagac aataaccctg ataaatgctt caataatatt gaaaaaggaa 9060  
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 9120  
tctgtttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag atcagttggg 9180  
tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg 9240  
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 9300  
atcccgtgtt gacgcggggc aagagcaact cggtcgccgc atacactatt ctcagaatga 9360  
cttgggttag tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga 9420  
attatgcagt gctgccataa ccactgagtg taacactcgc gccaaacttac ttctgacaac 9480  
gatcgaggga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 9540  
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 9600  
gatgcctgca gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 9660  
agcttcccgg caacaattaa tagactggat ggaggcggat aaagtgtcag gaccacttct 9720  
gcgctcggcc ctccggctg gctggtttat cactggggcc agatgggtaag tctggagccg gtgagcgtgg 9780  
gtctcggggt atcattgcag caactatgga tgaacgaaat agacagatcg ctgagatagg 9840  
ctacacgagc gggagttagc ggtaactgtc agaccaagtt tactcatata tacttttagat 9900  
tgctcactg attaagcatt aattttaaag gatctaggtg aagatccttt ttgataatct 10020  
tgatttaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 10080  
catgacaaaa tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaaaaa 10140  
gatcaaaagg taccagcgg tggtttgttt gccggatcaa gagctaccaaa ctctttttcc 10200  
aaaaccaccg ggcttcagca gaggcgagat accaaatact gtccttctag tgtagccgta 10260  
gaaggtaact cacttcaaga actctgtagc accgcctaca tacctcgctc tgtaattcct 10320  
gttaggccac gctgctgcca gtggcgataa gtcgtgtcct accgggttgg actcaagacg 10380  
gttaccagtg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 10440  
atagttaccg acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 10500  
cctggagcga caggggagaa aggggagctt cagggggaaa cgctgggtat ctttatagtc ctgtcgggtt 10560  
cacgcttccc gaaggagagc gctgctgcca gtgattttt gtgatgctcg tcaggggggc ggagcctatg 10620  
agagcgcacg agcaacgcgg ccttttttac gttcctggcc ttttgctggc cttttgctca 10680  
tgccacacct tgacttgagc gctcgccgca gccgaacgac tgtggataac cgtattaccg cctttgagt 10740  
gaaaaacgcc catgttcttt cctgcgttat cccctgattc cgagcgcagc gagtcaagtga gcgaggaagc 10800  
catgttcttt gctcgccgca attttctcct tacgcatctg tgcggtattt cacaccgcat 10860  
agctgatacc ctgatgcggg tctgctctga tgcgcgatag ttaagccagt atacactccg 10920  
ggaagagcgc ctgactgagg catggctgag ccccgacacc cgccaacacc cgctgacgag 11040  
atgggtgact gtgactgggt cccggcatcc gcttacagac aagctgtgac cgtctccggg 11100  
ctatcgctac ccttgacggg gtcagaggtt ttcaccgtca tcaccgaaac gcgagaggca gc 11152

&lt;210&gt; 13

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: primer

&lt;400&gt; 13

gacggatcgg gagatctcc

<210> 14  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 14  
ccgcctcaga agccatagag cc

22

<210> 15  
<211> 14455  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: plasmid

<400> 15  
aagcttgggc agaaatgggt gaactcccga gagtgtccta cacctagggg agaagcagcc 60  
aaggggttgt ttcccaccaa ggacgacccg tctgcgcaca aacggatgag cccatcagac 120  
aaagacatat tcattctctg ctgcaaaactt ggcataagctc tgctttgcct ggggctattg 180  
ggggaagttg cgggttcgtgc tcgcaggggt ctcacccttg actcttttaa tagctcttct 240  
gtgcaagatt acaatctaaa caattcggag aactcgacct tcctcctgag gcaaggacca 300  
cagccaactt cctcttacia gccgcacga ttttgcctt cagaaataga aataagaatg 360  
cttgctaaaa attatatttt taccaataag accaatccaa taggtagatt attagttact 420  
atggttaagaa atgaatcatt atcttttagt actattttta ctcaaattca gaagttagaa 480  
atgggaatag aaaatagaaa gagacgctca acctcaattg aagaacagggt gcaaggacta 540  
ttgaccacag gcctagaagt aaaaaaggga aaaaagagtg tttttgtcaa aataggagac 600  
aggtggtggc aaccagggac ttatagggga ccttacatct acagaccaac agatgcccc 660  
ttaccatata caggaagata tgacttaa atgggatagg ggggttacagt caatggctat 720  
aaagtgttat atagatccct cccttttctg gaaagactcg ccagagctag acctccttgg 780  
tgtatgttgt ctcaagaaga aaaagacgac atgaaacaac aggtacatga ttatatttat 840  
ctaggaacag gaatgcactt ttggggaaag attttccata ccaaggaggg gacagtggct 900  
ggactaatag aacattatct tgcaaaaact catggcatga gttattatga atagccttta 960  
ttggcccaac ctgcggttc ccagggtcta agtaagtttt tgggtacaaa ctgttcttaa 1020  
aacgaggatg tgagacaagt ggtttcctga cttgggttgg tatcaaagggt tctgatctga 1080  
gctctgagtg ttctattttc ctatgttctt ttggaattta tccaaatctt atgtaaatgc 1140  
ttatgtaaac caagatataa aagagtgtcg attttttgag taaacttgca acagtccctaa 1200  
cattcacctc ttgtgtgttt gtgtctgttc gccatcccgt ctccgctcgt cacttatcct 1260  
tcactttcca ggggtcccc ccgcagaccc cggcgaccct caggtcggcc gactgcggca 1320  
gctggcgccc gaacagggac cctcggataa gtgacccttg tctctatttc tactatttgg 1380  
tgtttgtctt gtattgtctc tttcttgtct ggctatcatc acaagagcgg aacggactca 1440  
ccatagggac caagctagcg actgaaaatg agacatatta tctgccacgg aggtgttatt 1500  
accgaagaaa tggccgccag tcttttggac cagctgatcg aagaggtact ggctgataat 1560  
cttcacctc ctagccattt tgaaccacct acccttcacg aactgtatga ttttagactg 1620  
acggcccccg aagatcccaa cgaggaggcg gtttcgcaga tttttcccga ctctgtaatg 1680  
ttggcggtgc aggaagggat tgacttactc acttttccgc cggcgcccgg ttctccggag 1740  
ccgcctcacc tttcccggca gcccgagcag ccggagcaga gagccttggg tccggtttct 1800  
atgccaaaac ttgtaccgga ggtgatcgat cttacctgcc acgaggctgg ctttccaccc 1860  
agtgcacgac aggatgaaga gggtagggag tttgtgttag attatgtgga gcaccccggg 1920  
cacggttgca ggtcttgtca ttatcaccgg aggaatacgg tattatgtgt 1980  
tcgcttttgt atatgaggac ctgtggcatg tttgtctaca gtaagtgaat attatgggca 2040  
gtgggtgata gagtgtgtgg tttgtgttgg taattttttt ttttaattttt acagttttgt 2100  
ggtttaaaaga attttgtatt gtgatttttt taaaagggtcc tctaccccgcg cctgacctga 2160  
agcccgagcc agaaccggag cctgcaagac ctaccggcga tagtagtacg gatagctgtg 2220  
tcctgagacg cccgacatca cctgtgtcta gagaatgcaa tagtagtacg ggtcccgtg 2280  
actccgggtc ttctaacaca cctcctgaga tacacccggg ggtcccgtg 2340  
aaccagttgc cgtgagagtt ggtgggcgtc gccaggctgt ggaatgtatc gaggacttgc 2400  
ttaacgagcc tgggcaacct ttggacttga gctgtaaacg cccagggcca taagggtgta 2460

|              |             |             |             |             |             |      |
|--------------|-------------|-------------|-------------|-------------|-------------|------|
| acctgtgatt   | gcgtgtgtgg  | ttaacgcctt  | tgtttgctga  | atgagttgat  | gtaagtttaa  | 2520 |
| ttaaaggggtga | gataatgttt  | aacttgcattg | gcgtgtttaa  | tggggcgggg  | cttaaagggg  | 2580 |
| atataatgcg   | ccgtgggcta  | atcttggtta  | catctgacct  | catggaggct  | tgggagtggt  | 2640 |
| tggaaagattt  | ttctgctgtg  | cgtaacttgc  | tggaacagag  | ctctaacagt  | acctcttggt  | 2700 |
| tttggagggtt  | tctgtggggc  | tcatcccagg  | caaagttagt  | ctgcagaatt  | aaggaggatt  | 2760 |
| acaagtggga   | atttgaagag  | cttttgaaat  | cctgtgtgga  | gctgtttgat  | tctttgaatc  | 2820 |
| tgggtcacca   | ggcgcttttc  | caagagaagg  | tcatcaagac  | tttggatttt  | tccacaccgg  | 2880 |
| ggcgcgctgc   | ggctgctgtt  | gcttttttga  | gttttataaa  | ggataaatgg  | agcgaagaaa  | 2940 |
| cccactctgag  | cgggggggtac | gtctgtggatt | ttctggccat  | gcatctgtgg  | agagcggttg  | 3000 |
| tgagacacaa   | gaatcgccctg | ctactgttgt  | cttccgtccg  | ccggcgata   | ataccgacgg  | 3060 |
| aggagcagca   | gcagcagcag  | gaggaagcca  | ggcgcgcgcg  | gcaggagcag  | agcccatgga  | 3120 |
| acccgagagc   | cggcctggac  | cctcggaat   | gaatgttgta  | cagggtggctg | aactgtatcc  | 3180 |
| agaactggaga  | gcatttttga  | caattacaga  | ggatgggcag  | gggctaaagg  | gggtaaagag  | 3240 |
| ggagcggggg   | gcttgtgagg  | ctacagagga  | ggctaggaat  | ctagctttta  | gcttaatgac  | 3300 |
| cagacaccgt   | cctgagtgta  | ttacttttca  | acagatcaag  | gataattgctg | ctaatgagct  | 3360 |
| tgtactgtctg  | gcgcagaagt  | attccataga  | gcagctgacc  | acttactggc  | tgcagccagg  | 3420 |
| ggatgattttt  | gaggaggcta  | ttagggata   | tgcaaagggtg | gcacttaggc  | cagattgcaa  | 3480 |
| gtacaagatc   | agcaaaacttg | taaataatcag | gaattgttgc  | tacatttctg  | ggaacggggc  | 3540 |
| cgagggtggag  | atagatacgg  | aggatagggt  | ggcctttaga  | tgtagcatga  | taaataatgtg | 3600 |
| gccgggggtg   | cttggcatgg  | acgggggtgg  | tattatgaat  | gtaagggttta | ctggcccaa   | 3660 |
| ttttagcggg   | acgggttttcc | tggccaatac  | caaccttatc  | ctacacgggtg | taagcttcta  | 3720 |
| tgggtttaac   | aatacctgtg  | tggaagcctg  | gaccgatgta  | aggggttcggg | gctgtgcctt  | 3780 |
| ttactgctgc   | tggaaggggg  | tggtgtgtcg  | ccccaaaagc  | agggcttcaa  | ttaagaaatg  | 3840 |
| cctctttgaa   | agggtgtacct | tggtgtatcct | gtctgagggt  | aactccaggg  | tgcgccacaa  | 3900 |
| tgtggcctcc   | gactgtgggt  | gcttcatgct  | agtgaagagc  | gtggctgtga  | ttaagcataa  | 3960 |
| catggatgtg   | ggcaactgcg  | aggacagggc  | ctctcagatg  | ctgacctgct  | cggacggcaa  | 4020 |
| ctgtcacctg   | ctgaagacca  | ttcacgtagc  | cagccactct  | cgcaaggcct  | ggccagtgtt  | 4080 |
| tgagcataac   | atactgaccc  | gctgttccct  | gcatttgggt  | aacaggaggg  | gggtgttccct | 4140 |
| accttaccaa   | tgcaatttga  | gtcacactaa  | gatattgctt  | gagcccgaga  | gcatgtccaa  | 4200 |
| gggtgaacctg  | aacgggggtgt | ttgacatgac  | catgaagatc  | tggaagggtgc | tgaggtagca  | 4260 |
| tgagaccgcg   | accaggtgca  | gaccctgcga  | gtgtggcggt  | aaacatatta  | ggaaccagcc  | 4320 |
| tgtgatgctg   | gatgtgaccg  | aggagctgag  | gcccgatcac  | ttgggtgctgg | cctgcaccgc  | 4380 |
| cgctgagttt   | ggctctagcg  | atgaagatac  | agattgaggt  | actgaaatgt  | gtggcgctgg  | 4440 |
| cttaagggtg   | ggaaagaata  | tataagggtg  | gggtcttatg  | tagttttgta  | tctgttttgc  | 4500 |
| agcagccgcc   | gcccgcctga  | gcaccaactc  | gtttgatgga  | agcattgtga  | gctcatattt  | 4560 |
| gacaacgcgc   | atgcccccat  | gggcgggggt  | gcgtcagaat  | gtgatgggct  | ccagcattga  | 4620 |
| tggtcgcccc   | gtcctgcccc  | caaactctac  | taccttgacc  | tacgagaccg  | tgctcggaac  | 4680 |
| gccgttgagg   | actgcagcct  | cgcgcgcgc   | ttcagccgct  | gcagccaccg  | cccgcgggat  | 4740 |
| tgtgactgac   | tttgctttcc  | tgagcccgct  | tgcaagcagt  | gcagcttccc  | ggtcatccgc  | 4800 |
| ccgcgatgac   | aagttgacgg  | ctcttttggc  | acaattggat  | tctttgaccc  | gggaacttaa  | 4860 |
| tgtcgtttct   | cagcagctgt  | tggtatctgc  | ccagcagggt  | tctgccctga  | aggcttctct  | 4920 |
| ccctcccaat   | gcggttttaa  | acataaataa  | aaaaccagac  | tctgtttgga  | tttggatcaa  | 4980 |
| gcaagtgtct   | tgctgtctct  | cgagggatct  | ttgtgaagga  | accttacttc  | tgtgggtgtga | 5040 |
| cataattgga   | caaactacct  | acagagattt  | aaagctctaa  | ggtaaatata  | aaatttttaa  | 5100 |
| gtgtataatg   | tggttaaacta | ctgattctaa  | ttgtttgtgt  | atttttagatt | ccaacctatg  | 5160 |
| gaactgatga   | atgggagcag  | tggtggaatg  | cctttaatga  | ggaaaacctg  | ttttgtctag  | 5220 |
| aagaaatgcc   | atctagtgat  | gatgaggcta  | ctgctgactc  | tcaacattct  | actcctccaa  | 5280 |
| aaaagaagag   | aaaggtagaa  | gaccccaagg  | actttccttc  | agaattgcta  | agttttttga  | 5340 |
| gtcatgctgt   | gtttagtaat  | agaactcttg  | ccttgctttgc | tatttacacc  | acaaaggaaa  | 5400 |
| aagctgcact   | gctatacaag  | aaaattatgg  | aaaaatatct  | tgtaaccttt  | ataagtaggc  | 5460 |
| ataacagtta   | taatcataac  | atactgtttt  | ttcttactcc  | acacaggcat  | agagtgtctg  | 5520 |
| ctattaataa   | ctatgctcaa  | aaattgtgta  | ccttttagctt | tttaatttgt  | aaaggggtta  | 5580 |
| ataaggaata   | tttgatgtat  | agtgccctga  | ctagagatca  | taatcagcca  | taccacattt  | 5640 |
| gtagagggttt  | tacttgcttt  | aaaaaacctc  | ccacacctcc  | ccctgaacct  | gaaacataaa  | 5700 |
| atgaatgcaa   | ttgttgttgt  | taacttggtt  | atttgcagctt | ataatgggtta | caataaagc   | 5760 |
| aatagcatca   | caaatttcac  | aaataaagca  | tttttttcac  | tgcattctag  | ttgtgggtttg | 5820 |
| tccaaactca   | tcaatgtatc  | ttatcatgtc  | tggatccggc  | tgtggaatgt  | gtgtcagtta  | 5880 |
| gggtgtggaa   | agtccccagg  | ctccccagca  | ggcagaagta  | tgcaaagcat  | gcatctcaat  | 5940 |
| tagtcagcaa   | ccaggtgtgg  | aaagtcccca  | ggctccccag  | caggcagaag  | tatgcaaagc  | 6000 |
| atgcatctca   | attagtcagc  | aaccatagtc  | ccgccccctaa | ctccgcccct  | cccgcgcccta | 6060 |
| actccgcccc   | gttccgcccc  | ttctccgccc  | catggctgac  | taattttttt  | tatttatgca  | 6120 |
| gaggccgagg   | ccgcctcgcc  | ctctgagcta  | ttccagaagt  | agtgaggagg  | cttttttgga  | 6180 |

|             |             |            |             |             |             |      |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| ggcctaggt   | tttgcaaaaa  | gcttggacac | aagacaggct  | tgcgagatat  | gtttgagaat  | 6240 |
| accactttat  | cccgcgtcag  | ggagaggcag | tgcgtaaaaa  | gacgcggact  | catgtgaaat  | 6300 |
| actggttttt  | agtgcgccag  | atctctataa | tctcgcgcaa  | cctattttcc  | cctcgaacac  | 6360 |
| tttttaagcc  | gtagataaac  | aggctgggac | acttcacatg  | agcgaataat  | acatcgtcac  | 6420 |
| ctgggacatg  | ttgcagatcc  | atgcacgtaa | actcgcaagc  | cgactgatgc  | cttctgaaca  | 6480 |
| atggaaaggc  | attattgccc  | taagccgtgg | cggtctggta  | ccgggtgcgt  | tactggcgcg  | 6540 |
| tgaactgggt  | attcgtcatg  | tcgataccgt | ttgtatttcc  | agctacgatc  | acgacaacca  | 6600 |
| gcgcgagctt  | aaagtgcgtg  | aacgcgcaga | aggcgatggc  | gaaggcttca  | tcgttattga  | 6660 |
| tgacctgggtg | gataccgggtg | gtactgcggg | tgcgattcgt  | gaaatgtatc  | caaaagcgca  | 6720 |
| ctttgtcacc  | atcttcgcaa  | aaccggctgg | tcgtcccgctg | gttgatgact  | atgttggtga  | 6780 |
| tatcccgcga  | gatacctgga  | ttgaacagcc | gtgggatatg  | ggcgctcgat  | tcgtcccggc  | 6840 |
| aatctccggt  | cgctaactct  | ttcaacgcct | ggcgctggcg  | ggcgctgttc  | tttttaactt  | 6900 |
| caggcgggtt  | acaatagttt  | ccagtaagta | ttctggaggc  | tgcatccatg  | acacaggcaa  | 6960 |
| acctgagcga  | aaccctgttc  | aaaccccgt  | ttaaaccatcc | tgaacccctc  | acgctagtcc  | 7020 |
| gccgctttaa  | tcacggcgca  | caaccgcctg | tcagtcggcg  | ccttgatggt  | aaaaccatcc  | 7080 |
| ctcactggta  | tcgcatgatt  | aaccgtctga | tgtggatctg  | gcgcggcatt  | gacccacgcg  | 7140 |
| aaatcctcga  | cgtccaggca  | cgtatttgtg | tgagcgatgc  | cgaacgtacc  | gacgatgatt  | 7200 |
| tatacgatac  | ggtgattggc  | taccgtggcg | gcaactggat  | ttatgagtgg  | gccccggatc  | 7260 |
| tttgtgaagg  | aaccttactt  | ctgtggtgtg | acataattgg  | acaaactacc  | tacagagatt  | 7320 |
| taaagctcta  | aggtaaatat  | aaaattttta | agtgtataat  | gtgttaaact  | actgattcta  | 7380 |
| attgtttgtg  | tatttttagat | tccaacctat | ggaactgatg  | aatgggagca  | gtggtggaat  | 7440 |
| gcctttaatg  | aggaaaacct  | gttttgctca | gaagaaatgc  | catctagtga  | tgtatgaggt  | 7500 |
| actgctgact  | ctcaacattc  | tactcctcca | aaaaagaaga  | gaaaggtaga  | agaccccaag  | 7560 |
| gactttcctt  | cagaattgct  | aagttttttg | agtcattgctg | tgtttagtaa  | tagaactcct  | 7620 |
| gcttgctttg  | ctattttacac | cacaaaggaa | aaagctgcac  | tgctatacaa  | gaaaattatg  | 7680 |
| gaaaaatatt  | ctgtaacctt  | tataagtagg | cataacagtt  | ataatcataa  | catactgttt  | 7740 |
| tttcttactc  | cacacaggca  | tagagtgtct | gctattaata  | actatgctca  | aaaatttgtt  | 7800 |
| acctttagct  | ttttaatttg  | taaaggggtt | aataaggaat  | atltgatgta  | tagtgccctg  | 7860 |
| actagagatc  | ataatcagcc  | ataccacatt | tgtagagggt  | ttacttgctt  | taaaaaacct  | 7920 |
| cccacacctc  | ccctgaacc   | tgaaacataa | aatgaatgca  | attgttgttg  | ttacttgctt  | 7980 |
| tattgtagct  | tataatgggt  | acaaataaag | caatagcatc  | acaaatttca  | caaataaagc  | 8040 |
| atltttttca  | ctgcattcta  | gttgtgtgtt | gtccaaactc  | atcaatgtat  | cttatcatgt  | 8100 |
| ctggatcccc  | aggaagctcc  | tctgtgtcct | cataaaccct  | aacctcctct  | acttgagagg  | 8160 |
| acattccaat  | cataggctgc  | ccatccaccc | tctgtgtcct  | cctgttaatt  | aggtcactta  | 8220 |
| acaaaaagga  | aattgggtag  | gggtttttca | cagaccgctt  | tctaagggtg  | atlttaaaat  | 8280 |
| atctgggaag  | tcccttccac  | tgctgtgttc | cagaagtgtt  | ggtaaacagc  | atlttaaaat  | 8340 |
| caacagcaga  | aacatacaag  | ctgtcagctt | tgcaacaagg  | cccaacaccc  | ccacaaatgt  | 8400 |
| gaagcactgt  | gggtgctgtg  | ttagtaatgt | gcaaaacagg  | aggcacattt  | tgctcatcaa  | 8460 |
| tgtagggtcc  | aaaatatcta  | gtgttttcat | ttttacttgg  | atcaggaacc  | tccccacctg  | 8520 |
| ctgggataagc | attatcctta  | tccaaaacag | ccttggtgtc  | agtggttcatc | cagcactcca  | 8580 |
| caactgtagc  | attttttggg  | gttacagttt | gagcaggata  | tttggtcctg  | tgctgactgt  | 8640 |
| acacaccctg  | cagctccaaa  | ggttccccac | caacagcaaa  | aaaatgaaaa  | tttgaccctt  | 8700 |
| gaatgggttt  | tccagcacca  | ttttcatgag | ttttttgtgt  | ccctgaatgc  | aagtttaaca  | 8760 |
| tagcagttac  | cccaataacc  | tcagttttta | cagtaacagc  | ttccacatc   | aaaatatatt  | 8820 |
| cacagggttaa | gtcctcattt  | aaattaggca | aaggaattct  | tgaagacgaa  | agggcctcgt  | 8880 |
| gatacgcccta | tttttatagg  | ttaatgtcat | gataataatg  | gtttctttaga | cgtcagggtg  | 8940 |
| cacttttccg  | ggaaatgtgc  | gcggaacccc | tatttggtta  | tttttctaaa  | tacattcaaa  | 9000 |
| tatgtatccg  | atcaacattt  | aataaccctg | ataaatgctt  | caataatatt  | gaaaaaggaa  | 9060 |
| gagtatgagt  | gtcaccacag  | tccgtgtcgc | ccttattccc  | ttttttgcgg  | cattttgcct  | 9120 |
| tcctgttttt  | ggttacatcg  | aaacgctggg | gaaagtaaaa  | gatgtcgaag  | atcagttggg  | 9180 |
| tgcacgagtg  | cggtttccaa  | aactggatct | caacagcggg  | aagatccttg  | agagttttcg  | 9240 |
| ccccgaagaa  | gacgccccgg  | tgatgagcac | tttttaaagt  | ctgctatgtg  | gcgcgggtatt | 9300 |
| atcccgtgtt  | tactcaccag  | aagagcaact | cggtcgcggc  | atacactatt  | ctcagaatga  | 9360 |
| cttggttgag  | gctgccataa  | tcacagaaaa | gcactcttac  | gatggcatga  | cagtaagaga  | 9420 |
| attatgcagt  | ccatgagtga  | ccatgagtga | taacactgcg  | gccaaacttac | ttctgacaac  | 9480 |
| gatcggagga  | ccgaaggagc  | taaccgcttt | tttgacaaac  | atgggggatc  | atgtaactcg  | 9540 |
| ccttgatcgt  | tggaacccgg  | agctgaatga | agccatacca  | aacgacgagc  | gtgacaccac  | 9600 |
| gatgcttcga  | gcaatggcaa  | caacgttgcg | caaactatta  | actggcgaac  | tacttactct  | 9660 |
| agcttcccgg  | caacaattaa  | tagactggat | ggaggcggat  | aaagtgtcag  | gaccacttct  | 9720 |
| gcgctcggcc  | cttcgggctg  | gctgggttat | tgctgataaa  | tctggagccg  | gtgagcgtgg  | 9780 |
| gtctcgcggg  | atcattgcag  | cactggggcc | agatggttaag | ccctcccgtg  | tcgtagtatt  | 9840 |
| ctacacgacg  | gggagtcagg  | caactatgga | tgaacgaaat  | agacagatcg  | ctgagatagg  | 9900 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tgccctactg  | attaagcatt  | ggtaactgtc  | agaccaagtt  | tactcatata  | tacttttagat | 9960  |
| tgatttaaaa  | cttcattttt  | aattttaaag  | gatctagggtg | aagatccctt  | ttgataatct  | 10020 |
| catgaccaaaa | atcccttaac  | gtgagttttc  | gttccactga  | gcgtcagacc  | ccgtagaaaa  | 10080 |
| gatcaaagga  | tcttcttgag  | atcctttttt  | tctgcgcgta  | atctgctgct  | tgcaaacaaa  | 10140 |
| aaaaccacccg | ctaccagcgg  | tggtttgttt  | gccggatcaa  | gagctaccaa  | ctctttttcc  | 10200 |
| gaaggtaact  | ggcttcagca  | gagcgcagat  | accaaatact  | gtccttctag  | tgtagccgta  | 10260 |
| gttaggccac  | cacttcaaga  | actctgtage  | accgcctaca  | tacctcgctc  | tgctaactct  | 10320 |
| gttacagtg   | gctgctgcc   | gtggcgataa  | gtcgtgtctt  | accgggttgg  | actcaagacg  | 10380 |
| atagttaccg  | gataaggcgc  | agcggtcggg  | ctgaacgggg  | ggttcgtgca  | cacagcccg   | 10440 |
| cttgagcgga  | acgacctaca  | ccgaactgag  | atacctacag  | cgtgagctat  | gagaaagcgc  | 10500 |
| cacgcttccc  | gaagggagaa  | aggcggacag  | gtatccggta  | agcggcaggg  | tcggaacagg  | 10560 |
| agagcgcaacg | agggagcttc  | cagggggaaa  | cgctgtgtat  | ctttatagtc  | ctgtcgggtt  | 10620 |
| tcgccacctc  | tgacttgagc  | gtcgattttt  | gtgatgctcg  | tcaggggggc  | ggagccctatg | 10680 |
| gaaaaacgcc  | agcaacgcgg  | ccttttttacg | gttcctggcc  | ttttgctggc  | cttttgctca  | 10740 |
| catgttcttt  | cctgcgttat  | cccctgattc  | tgtagataac  | cgtattaccg  | cctttgagtg  | 10800 |
| agctgatacc  | gctgcgcgca  | gccgaacgac  | cgagcgcagc  | gagtcagtga  | gcgaggaagc  | 10860 |
| ggaagagcgc  | ctgatgcggg  | attttctcct  | tacgcatctg  | tgccgtatct  | cacaccgcct  | 10920 |
| accgcctcag  | aagccataga  | gcccaccgca  | ccccagcat   | gcctgctatt  | gtcttcccaa  | 10980 |
| tcctccccct  | tgctgtcctg  | ccccacccca  | ccccacagaa  | tagaatgaca  | cctactcaga  | 11040 |
| caatgcgatg  | caatttcctc  | attttattag  | gaaaggacag  | tgaggagtggc | accttccagg  | 11100 |
| gtcaagggaag | gcacggggga  | ggggcaaaaca | acagatgggt  | ggcaactaga  | aggcacagtc  | 11160 |
| gaggctgatc  | agcgagctct  | agcatttagg  | tgactactata | gaataggggc  | ctctagatgc  | 11220 |
| atgctcgagc  | ggccgcttct  | ttattcttgg  | gcaatgtatg  | aaaaagtgt   | agaggatgtg  | 11280 |
| gcaaatatct  | cattaatgta  | gttgtggcca  | gaccagtccc  | atgaaaatga  | catagagtat  | 11340 |
| gcacttgag   | ttgtgtctcc  | tgtttctctg  | gtaccgttta  | gtgtaatggt  | tagtgttaca  | 11400 |
| ggtttagttt  | tgtctccgtt  | taagtaaaact | tgtagtaca   | tgtagtactt  | ggcagtttta  | 11460 |
| ccgtagagatt | ttggataagc  | tgtagggtta  | ggcataaaatc | caacagcgtt  | tgtagaggct  | 11520 |
| gtgccttcag  | taagatctcc  | atttctaaag  | ttccaatatt  | ctgggtccag  | gaagggaattg | 11580 |
| tttagtagca  | ctccattttc  | gtcaaatctt  | ataataagat  | gagcactttg  | aactgttcca  | 11640 |
| gataattggg  | ccaaactgcc  | tttaacagcc  | aaaactgaaa  | ctgtagcaag  | tatttgactg  | 11700 |
| ccacattttg  | ttaagaccaa  | agttaggtta  | gcattcttct  | ctgcatttag  | tctacagtta  | 11760 |
| ggagatggag  | ctgggtgtgg  | ccacaaagtt  | agcttatcat  | tatttttgtt  | tcctactgta  | 11820 |
| atggcacctg  | tgctgtcaaa  | actaaggcca  | tttccctagt  | taggaaccat  | agccttgttt  | 11880 |
| gaatcaaatg  | ctaggccatg  | gccaaatttt  | gttttgaggg  | gatttgtgtt  | tggtgcatta  | 11940 |
| ggtgaaccaa  | attcaagccc  | atctcctgca  | ttaatggcta  | tggtgttagc  | gtcaaacatc  | 12000 |
| aaccccttgg  | cagtgccttag | gttaacctca  | agcttttttg  | aattgtttga  | agctgtaaac  | 12060 |
| aagtaaaagg  | ctttgttgta  | gttaatatcc  | aagtgtgtgg  | ctgagtttat  | aaaaagaggg  | 12120 |
| ccctgtccta  | gtcttagatt  | tagttggttt  | tgagcatcaa  | acggataact  | aacatcaagt  | 12180 |
| ataaggcgtc  | tgttttgaga  | atcaatcctt  | agtcctcctg  | ctacattaag  | ttgcataattg | 12240 |
| ccttggtgaat | caaaaaccaa  | ggctccagta  | actttagttt  | gcaagggaag  | attattaata  | 12300 |
| gtcacacctg  | gaccagttgc  | tacggtcaaa  | gtgttttagt  | cgtctgttac  | atgcaaagga  | 12360 |
| gccccgtact  | ttagtccctag | ttttccattt  | tgtgtataaa  | tggtgtcttt  | caagtcaatg  | 12420 |
| cccaagctac  | cagtggcagt  | agtttagagg  | ggtagggcag  | tgatagtaag  | ggtagtcta   | 12480 |
| tcgggtgggtg | tgagggggcc  | tgatgtttgc  | agggctagct  | ttccttctga  | cactgtgagg  | 12540 |
| ggctccttggg | tgccaatgct  | aagtttgagg  | tcgtgcacgg  | ttagcggggc  | ctgtgattgc  | 12600 |
| atgggtgagtg | tgttgcccg   | gaccattaga  | gggtgcggcg  | cagccacagt  | tagggcttct  | 12660 |
| gaggtaaactg | tgaggggtgc  | agatatttcc  | aggtttatgt  | ttgacttggt  | ttttttgaga  | 12720 |
| gggtgggtca  | cagtgggttac | attttgggag  | gtaagggtgc  | cggcctcgct  | cagagagagg  | 12780 |
| ccgttgccca  | ttttgagcgc  | aagcatgcc   | ttggaggtaa  | ctagagggtc  | ggtagggcgc  | 12840 |
| aaagagagta  | ccccaggggg  | actctcttga  | aaccatttgg  | gggatacaaa  | gggaggagta  | 12900 |
| agaaaaggca  | cagttggagg  | accggtttcc  | gtgtcatatg  | gatacacggg  | gttgaaggta  | 12960 |
| tcttcagacg  | gtcttgccgc  | cttcatcttg  | gatctcaagc  | ctgccacacc  | tcacctcgac  | 13020 |
| catccgccgt  | ctcaagaccg  | cctactttaa  | ttacatcatc  | agcagcacct  | ccgccagaaa  | 13080 |
| caaccccgac  | cgccacccgc  | tgccgcccgc  | cacgggtgctc | agcctacctt  | gcgactgtga  | 13140 |
| ctgggttagac | gcctttctcg  | agaggttttc  | cgatccgggtc | gatgcggact  | cgctcaggtc  | 13200 |
| cctcggtggc  | ggagtaccgt  | tcggaggccg  | acgggtttcc  | gatccaagag  | tactggaaag  | 13260 |
| accgcgaaga  | gtttgtcctc  | aaccgcgagc  | ccaacagcga  | gctcgaattc  | agatccgagc  | 13320 |
| tcggtaccaa  | gcttggggtc  | ccctatagtg  | agtcgtatta  | atttcgataa  | gccagtaagc  | 13380 |
| agtgggttct  | ctagttagcc  | gggcggagtt  | gcttatatag  | acctcccacc  | gtacacgcct  | 13440 |
| accgcccatt  | tcggtcaatg  | cattgacgtc  | gttacgacat  | tttggaaggt  | cccgttgatt  | 13500 |
| ttggtgccaa  | aacaaactcc  | cattgatgta  | aatgggggtg  | agacttgaa   | atccccgtga  | 13560 |
| gtcaaacccg  | tatccacgcc  |             | ctgccaaaac  | cgcatcacca  | tggtaatagc  | 13620 |

|             |            |            |            |            |             |       |
|-------------|------------|------------|------------|------------|-------------|-------|
| gatgactaat  | acgtagatgt | actgccaa   | aggaaagtcc | cataaggtca | tgtactgggc  | 13680 |
| ataatgccag  | gcgggccatt | taccgtcatt | gacgtcaata | gggggcgtac | ttggcatatg  | 13740 |
| atacacttga  | tgtactgcca | agtgggcagt | ttaccgtaaa | tagtccaccc | attgacgtca  | 13800 |
| atggaaaagtc | cctattggcg | ttactatggg | aacatacgtc | attattgacg | tcaatgggcg  | 13860 |
| ggggtcggttg | ggcggtcagc | caggcgggcc | atttaccgta | agttatgtaa | cgcggaactc  | 13920 |
| catatatggg  | ctatgaacta | atgaccccg  | aattgattac | tattaataac | tagtcaataa  | 13980 |
| tcaatgtcaa  | cgcgtatata | tggcccgta  | atcggaagc  | agcgcaaaac | gcctaaccct  | 14040 |
| aagcagattc  | ttcatgcaat | tgtcggtcaa | gccttgctt  | gtttagctt  | aaattttgct  | 14100 |
| cgcgcactac  | tcagcgacct | ccaacacaca | agcaggagc  | agatactggc | ttaaactatgc | 14160 |
| ggcatcagag  | cagattgtac | tgagagtcga | ccatagggga | tcgggagatc | tcccgatccg  | 14220 |
| tctatgggtgc | actctcagta | caatctgctc | tgatgccgca | tagttaagcc | agatacact   | 14280 |
| ccgctatcgc  | tacgtgactg | ggcatggct  | gcgccccgac | acccgccaac | acccgctgac  | 14340 |
| gcgccttgac  | gggcttgtct | gctcccgca  | tccgcttaca | gacaagctgt | gaccgtctcc  | 14400 |
| gggagctgca  | tgtgtcagag | gttttcaccg | tcataccgca | aacgcgcgag | gcagc       | 14455 |

&lt;210&gt; 16

&lt;211&gt; 10610

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid

&lt;400&gt; 16

|             |             |             |            |             |             |      |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| gacggatcgg  | gagatccg    | cggtacacag  | aattcaggag | acacaactcc  | aagtgcatac  | 60   |
| tctatgtcat  | tttcatggga  | ctggctctggc | cacaactaca | ttaatgaaat  | atttgccaca  | 120  |
| tcctcttaca  | ctttttcata  | cattgcccac  | gaataaagaa | tcgtttgtgt  | tatgtttcaa  | 180  |
| cgtgtttatt  | tttcaattgc  | agaaaatttc  | aagtcatttt | tcattcagta  | gtatagcccc  | 240  |
| accaccacat  | agcttatata  | gatcacgta   | ccttaatcaa | actcacagaa  | ccctagtatt  | 300  |
| caacctgcca  | cctccctccc  | aacacacaga  | gtacacagtc | ctttctcccc  | ggctggcctt  | 360  |
| aaaaagcatc  | atatcatggg  | taacagacat  | attcttaggt | gttatattcc  | acacggtttc  | 420  |
| ctgtcgagcc  | aaacgctcat  | cagtcatatt  | aataaactcc | ccgggcagct  | cacttaagtt  | 480  |
| catgtcgtcg  | tccagctgct  | gagccacagg  | ctgctgtcca | acttgcgggt  | gcttaacggg  | 540  |
| cggcgaagga  | gaagtccacg  | cctacatggg  | ggtagagtca | taatcgtgca  | tcaggatagg  | 600  |
| gcgggtgggtc | tgcagcagcg  | cgccaataaa  | ctgctgccgc | cgccgctccg  | tcctgcagga  | 660  |
| atacaacatg  | gcagtgggtc  | cctcagcgat  | gattcgacc  | gcccgagca   | taaggcgct   | 720  |
| tgctctccgg  | gcacagcagc  | gcaccctgat  | ctcacttaaa | tcagcacagt  | aactgcagca  | 780  |
| cagcaccaca  | atattgttca  | aaatcccaca  | gtgcaaggcg | ctgtatccaa  | agctcatggc  | 840  |
| ggggaccaca  | gaaccacagt  | ggccatcata  | ccacaagcgc | aggtagatta  | agtggcgacc  | 900  |
| cctcataaac  | acgctgggaa  | taaacattac  | ctcttttggc | atgttgtaat  | tcaccacctc  | 960  |
| ccggtaccat  | ataaacctct  | gattaaacat  | ggcgccatcc | accaccatcc  | taaaccagct  | 1020 |
| ggccaaaacc  | tgcccgccgg  | ctatacactg  | cagggaaccg | ggactggaac  | aatgacagtg  | 1080 |
| gagagccag   | gactcgtaac  | catggatcat  | catgctcgtc | atgatataca  | tgttggcaca  | 1140 |
| acacaggcac  | acgtgcatac  | acttcctcag  | gattacaagc | tcctcccgcg  | ttagaaccat  | 1200 |
| atcccaggga  | acaacccatt  | cctgaatcag  | cgtaaatccc | acactgcagg  | gaagacctcg  | 1260 |
| cacgtaactc  | acgttgtgca  | ttgtcaaagt  | gttacattcg | ggcagcagcg  | gatgatcctc  | 1320 |
| cagtatggta  | gcgcgggttt  | ctgtctcaaa  | aggaggtaga | cgatccctac  | tgtacggagt  | 1380 |
| gcgcgagac   | aaccgagatc  | gtgttggtcg  | tagtgtcatg | ccaaatggaa  | cgccggacgt  | 1440 |
| agtcataatt  | cctgaagcaa  | aaccagggtc  | gggcgtgaca | aacagatctg  | cgtctccggt  | 1500 |
| ctgcgcgctt  | agatcgctct  | gtgtagtagt  | tgtagtatat | ccactctctc  | aaagcatcca  | 1560 |
| ggcgccccct  | ggcttcgggt  | tctatgtaaa  | ctccttcatg | cgccgctgcc  | ctgataacat  | 1620 |
| ccaccaccgc  | agaataagcc  | acaccagcc   | aacctacaca | ttcgttctgc  | gagtcacaca  | 1680 |
| cgggaggagc  | gggaagagct  | ggaagaacca  | tgtttttttt | tttattccaa  | aagattatcc  | 1740 |
| aaaacctcaa  | aatgaagatc  | tattaagtga  | acgcgctccc | ctccggtggc  | gtgggtcaaac | 1800 |
| tctacagcca  | aagaacagat  | aatggcattt  | gtaagatgtt | gcacaatggc  | ttccaaaagg  | 1860 |
| caaacggccc  | tcacgtccaa  | gtggacgtaa  | aggctaaacc | cttcagggtg  | aatctcctct  | 1920 |
| ataaacattc  | cagcaccttc  | aacctgccc   | aaataattct | catctcgcca  | ccttctcaat  | 1980 |
| atatctctaa  | gcaaatcccg  | aattattaagt | ccggccattg | taaaaatctg  | ctccagagcg  | 2040 |
| ccctccacct  | tcagcctcaa  | gcagcgaatc  | atgattgcaa | aaattcagggt | tcctcacaga  | 2100 |
| cctgtataag  | attcaaaaagc | ggaacattaa  | caaaaatacc | gcgatcccg   | aggtcccttc  | 2160 |
| gcaggggccag | ctgaacataa  | tcgtgcagggt | ctgcacggac | cagcgcgggc  | acttcccgcg  | 2220 |



|              |             |             |             |             |              |      |
|--------------|-------------|-------------|-------------|-------------|--------------|------|
| caggaacctt   | gacaaaagaa  | cccacactga  | ttatgacacg  | catactcgga  | gctatgctaa   | 2280 |
| ccagcgtagc   | cccgatgtaa  | gctttgtgtg  | atgggcggcg  | atataaaatg  | caaggtgctg   | 2340 |
| ctcaaaaaat   | caggcaaaagc | ctcgcgcaaa  | aaagaaaagca | catcgtagtc  | atgctcatgc   | 2400 |
| agataaaaggc  | aggtaagctc  | cggaaaccacc | acagaaaaag  | acaccatttt  | tctctcaaac   | 2460 |
| atgtctgcgg   | gtttctgcat  | aaacacaaaa  | taaaataaca  | aaaaaacatt  | taaacattag   | 2520 |
| aagcctgtct   | tacaacagga  | aaaacaaccc  | ttataagcat  | aagacggact  | acggccatgc   | 2580 |
| cggcgtgacc   | gtaaaaaaac  | tggtcaccgt  | gattaaaaag  | caccaccgac  | agctcctcgg   | 2640 |
| tcatgtccgg   | agtacataatg | taagactcgg  | taaacacatc  | aggttgattc  | atcggtcagt   | 2700 |
| gctaaaaaagc  | gaccgaaata  | gcccggggga  | atacatacc   | gcaggcgtag  | agacaacatt   | 2760 |
| acagccccc    | taggaggtat  | aacaaaatta  | ataggagaga  | aaaacacata  | aacacctgaa   | 2820 |
| aaaccctcct   | gcctaggcaa  | aatagcacc   | tcccgcctca  | gaacaacata  | cagcgcttca   | 2880 |
| cagcggcagc   | ctaacagtca  | gccttaccag  | taaaaaagaa  | aacctattaa  | aaaaacacca   | 2940 |
| ctcgacacgg   | caccagctca  | atcagtcaca  | gtgtaaaaaa  | gggccaagtg  | cagagcgagt   | 3000 |
| atataatagga  | ctaaaaaatg  | acgtaacggt  | taaaagtcac  | aaaaaacacc  | caataaacccg  | 3060 |
| cacgcgaacc   | tacgcccaga  | aacgaaaagc  | aaaaaaccca  | caacttcctc  | agaactcgtcac | 3120 |
| ttcgttttc    | ccacgttacg  | taacttcctc  | gacccctccc  | cgatccccta  | tggtcgactc   | 3180 |
| tcagtacaat   | ctgctctgat  | gccgcatagt  | taagccagta  | tctgctccct  | gcttggtgtg   | 3240 |
| tgagggtcgc   | tgagttagtc  | gcgagcaaaa  | tttaagctac  | aacaaggcaa  | ggcttgaccg   | 3300 |
| acaattgcat   | gaagaatctg  | cttagggtta  | ggcgttttgc  | gctgcttcgc  | gatgtacggg   | 3360 |
| ccagatatatac | gcgttgacat  | tgattattga  | ctagttatta  | atagtaatca  | attacggggg   | 3420 |
| cattagttca   | tagcccatat  | atggagttcc  | gcgttacata  | acttacggta  | aatggcccg    | 3480 |
| ctggctgacc   | gcccacacgac | ccccgccc    | tgacgtcaat  | aatgacgtat  | gttcccatag   | 3540 |
| taacggcaat   | agggactttc  | cattgacgtc  | aatgggtgga  | ctatttacgg  | taaactgccc   | 3600 |
| acttgccagt   | acatcaagt   | tatcatatgc  | caagtacgcc  | ccctattgac  | gtcaatgacg   | 3660 |
| gtaaatggcc   | cgccctggcat | tatgcccagt  | acatgacctt  | atgggacctt  | cctacttggc   | 3720 |
| agtacatcta   | cgtattagtc  | atcgctatta  | ccatgggtgat | gcgggttttg  | cagtacatca   | 3780 |
| atgggcgtgg   | gactcacggg  | gactcacggg  | gattttccaa  | tctccacccc  | attgacgtca   | 3840 |
| atgggagttt   | gttttggcac  | caaaatcaac  | gggactttcc  | aaaatgtcgt  | aacaactccg   | 3900 |
| ccccattgac   | gcaaattggg  | ggtaggcgtg  | tacggtggga  | ggtctatata  | agcagagctc   | 3960 |
| tctggctaac   | tagagaaccc  | actgcttact  | ggcttatcga  | aattaatacg  | actcactata   | 4020 |
| gggagaccca   | agcttggtac  | cgagctcgga  | tctgaattcg  | agctcgctgt  | tgggctcgcg   | 4080 |
| gttgaggaca   | aactcttcgc  | ggtctttcca  | gtactcttgg  | atcggaacc   | cgtcgccctc   | 4140 |
| cgaacgggtac  | tccgccaccg  | agggacactga | gcgagtccgc  | atcgaccgga  | tcggaaaacc   | 4200 |
| tctcgagaaa   | ggcgtctaac  | cagtcacagt  | cgcaaggtag  | gctgagcacc  | gtggcggg     | 4260 |
| gcagcggtg    | gcggctcggg  | ttgtttctgg  | cggaggtgct  | gctgatgatg  | taattaaagt   | 4320 |
| aggcggtctt   | gagacggcgg  | atggctcgagg | tgaggtgtgg  | caggcttgag  | atccaagatg   | 4380 |
| aagcgcgcaa   | gaccgtctga  | agataccctc  | aaccccggtg  | atccatata   | cacggaaacc   | 4440 |
| ggtcctccaa   | ctgtgccttt  | tcttactcct  | ccctttgtat  | cccccaatgg  | gtttcaagag   | 4500 |
| agtcctccctg  | gggtactctc  | tttgcgccta  | tccgaacctc  | tagttacctc  | caatggcatg   | 4560 |
| cttgcgctca   | aaatgggcaa  | cggcctctct  | ctggacgagg  | ccggcaacct  | tacctcccaa   | 4620 |
| aatgtaacca   | ctgtgagccc  | acctctcaaa  | aaaaccaagt  | caaacataaa  | cctggaaata   | 4680 |
| tctgcacccc   | tcacagttac  | ctcagaagcc  | ctaactgtgg  | ctgcgcgcgc  | acctctaata   | 4740 |
| gtcgcgggca   | acacactcac  | catgcaatca  | caggccccgc  | taaccgtgca  | cgactccaaa   | 4800 |
| cttagcattg   | ccacccaagg  | acccctcaca  | gtgtcagaag  | gaaagctagc  | cctgcaaaaa   | 4860 |
| tcaggccccc   | tcaccaccac  | cgatagcagt  | acccttacta  | tcactgcctc  | acccctcta    | 4920 |
| actactgcc    | ctggtagctt  | gggcattgac  | ttgaaagagc  | ccattttatac | acaaaaatgga  | 4980 |
| aaactaggac   | taaagtacgg  | ggctcctttg  | catgtaacag  | acgacctaaa  | cactttgacc   | 5040 |
| gtagcaactg   | gtccaggtgt  | gactattaat  | aatacttctt  | tgcaaaactaa | agttactgga   | 5100 |
| gccttgggtt   | ttgattcaca  | aggcaatatg  | caacttaatg  | tagcaggagg  | actaaggatt   | 5160 |
| gatttctcaaa  | acagacgcct  | tatacttgat  | gttagttatc  | cgtttgatgc  | tcaaaaccaa   | 5220 |
| ctaaatctaa   | gactaggaca  | gggcccctct  | tttataaaact | cagcccacaa  | cttggatatt   | 5280 |
| aactacaaca   | aaggcccttta | cttgtttaca  | gcttcaaaaa  | attccaaaaa  | gcttgagggt   | 5340 |
| aacctaagca   | ctgccaaggg  | gttgatgttt  | gacgctacag  | ccatagccat  | taatgcagga   | 5400 |
| gatgggcttg   | aatttggttc  | acctaataga  | ccaaacacaa  | atcccctcaa  | aacaaaaatt   | 5460 |
| ggccatggcc   | tagaatttga  | gttatggttc  | gctatggttc  | ctaaactagg  | aactggcctt   | 5520 |
| agttttgaca   | gcacaggtgc  | cattacagta  | ggaaacaaaa  | ataatgataa  | gctaactttg   | 5580 |
| tggaaccacac  | cagctccatc  | tcctaactgt  | agactaaatg  | cagagaaaaga | tgctaaactc   | 5640 |
| actttgggtc   | taacaaaaatg | tggcagtcac  | atacttgcta  | cagtttcagt  | tttggtgtgt   | 5700 |
| aaaggcagtt   | tggtccaat   | atctggaaca  | gttcaaaagt  | ctcatcttat  | tataagattt   | 5760 |
| gacgaaaaatg  | gagtgctact  | aaacaattcc  | ttcctggacc  | cagaatattg  | gaactttaga   | 5820 |
| aatggagatc   | ttactgaagg  | cacagcctat  | acaaacgctg  | ttggatttat  | gcctaacccta  | 5880 |
| tcagcttatc   | caaaaatctca | cggtaaaact  | gccaaaaagta | acattgtcag  | tcaagtttac   | 5940 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ttaaaccggag | acaaaactaa  | acctgtaaca  | ctaaccatta  | cactaaacgg  | tacacaggaa  | 6000 |
| acaggagaca  | caactccaag  | tgcatactct  | atgtcatttt  | catgggactg  | gtctggccac  | 6060 |
| aactacatta  | atgaaatatt  | tgccacatcc  | tcttacactt  | tttcatacat  | tgcccaagaa  | 6120 |
| taaagaagcg  | gccgctcgag  | catgcatcta  | gagggcccta  | ttctatagt   | tcacctaaat  | 6180 |
| gctagagctc  | gctgatcagc  | ctcgactgtg  | ccttctagtt  | gccagccatc  | tgttggtttg  | 6240 |
| ccctcccccg  | tgcttctctt  | gaccctggaa  | ggtgccactc  | ccactgtcct  | ttcctaataa  | 6300 |
| aatgaggaaa  | ttgcatcgca  | ttgtctgagt  | aggtgtcatt  | ctattctggg  | gggtgggggtg | 6360 |
| gggcaggaca  | gcaaggggga  | ggattgggaa  | gacaatagca  | ggcatgctgg  | ggatgcgggtg | 6420 |
| ggctctatgg  | cttctgaggg  | ggaaagaacc  | agctggggct  | ctagggggta  | tccccacggc  | 6480 |
| ccctgtagcg  | gcgcatthaag | cgcgggcggt  | gtgggtggta  | cgcgagcggt  | gaccgctaca  | 6540 |
| cttgccagcg  | ccctagcgcc  | cgctccttcc  | gcttcttccc  | cttctcttct  | cgccacgttc  | 6600 |
| gcccgtcttc  | cccgcaagc   | tctaaatcgg  | ggcatccctt  | tagggttccg  | atctagtgtc  | 6660 |
| ttacggcacc  | tgcaccccaa  | aaaacttgat  | taggggtgat  | gttcacgtag  | tggggccatcg | 6720 |
| ccctgataga  | cggttttttcg | ccctttgacg  | ttggagtcca  | cgttctttaa  | tagtggtactc | 6780 |
| ttgttccaaa  | ctggaacaac  | actcaaccct  | atctcggtct  | attcttttga  | tttataaggg  | 6840 |
| attttgggga  | tttcggccta  | ttggttaaaa  | aatgagctga  | tttaacaaaa  | atttaacgcg  | 6900 |
| aattaattct  | gtggaatgtg  | tgctcagttag | gggtgtggaaa | gtccccaggc  | tccccaggca  | 6960 |
| ggcagaagta  | tgcaaaagcat | gcatctcaat  | tagtcagcaa  | ccagggtgtg  | aaagtcccca  | 7020 |
| ggctccccag  | caggcagaag  | cccgcccccta | atgcatctca  | attagtgcag  | aaccatagtc  | 7080 |
| cgccccctaa  | ctccgccccat | tatttatgca  | actccgcccc  | gttccgcccc  | ttctccgccc  | 7140 |
| catggctgac  | taattttttt  | tatttatgca  | gagggccgagg | ccgctctg    | ctctgagcta  | 7200 |
| ttccagaagt  | agtgaggagg  | atctgatcaa  | ggcctaaggct | tttgcaaaaa  | gctcccgga   | 7260 |
| gcttgatat   | ccatttttcg  | agggttctccg | gagacaggat  | gaggatcggt  | tcgcatgatt  | 7320 |
| gaacaagatg  | gattgcacgc  | cggtctgctc  | gcccgttggg  | tggagaggct  | attcggctat  | 7380 |
| gactgggcac  | aacagacaat  | caagaccgac  | gatgcgcgcg  | tggtccggct  | gtcagcgag   | 7440 |
| gggcgccccg  | ttctttttgt  | gctggccacg  | ctgtccgggtg | ccctgaatga  | actgcaggac  | 7500 |
| gaggcagcgc  | ggctatcggt  | ggactggctg  | acgggcgttc  | cttgcgcagc  | tgtgctcgac  | 7560 |
| gttgtcactg  | aagcgggaag  | tgccgagaaa  | ctattggggc  | aagtgccggg  | gcaggatctc  | 7620 |
| ctgtcatctc  | accttgctcc  | tacctgcccc  | gtatccatca  | tggctgatgc  | aatgcggcgg  | 7680 |
| ctgcatacgc  | ctcgcatgga  | agccgggtctt | ttcgaccacc  | aagcgaaaca  | tcgcatcgag  | 7740 |
| cgagcacgta  | cgccagccga  | actggttcgcc | gtcgatcagg  | atgatctgga  | cgaagagcat  | 7800 |
| caggggctcg  | tgacccatgg  | cgatgcctgc  | aggctcaagg  | cgcgcatgcc  | cgacggcgag  | 7860 |
| gatctcgtcg  | tcacgcactg  | tgccgggctg  | ttgcccgaata | tcattggtgga | aaatggccgc  | 7920 |
| ttttctggat  | gtgatattgc  | tggaagagctt | gggtgtggcg  | accgctatca  | ggacatagcg  | 7980 |
| ttggctaccc  | tcgcccgtcc  | cgattcgcag  | ggcggcgaat  | gggctgaccg  | cttccctcgtg | 8040 |
| ctttacggta  | cgggactctg  | gggttcgaaa  | cgcatcgcc   | tctatcgcc   | ctttgacgag  | 8100 |
| ttcttctgag  | cgattccacc  | gcccgccttct | tgaccgacca  | agcgacgccc  | aacctgccat  | 8160 |
| cacgagattt  | ctggatgatc  | ctccagcgcg  | atgaaagggtt | gggcttcgga  | atcggttttcc | 8220 |
| gggacgcccg  | tattgcagct  | tataatgggt  | gggatctcat  | gctggagttc  | ttcgcccacc  | 8280 |
| ccaacttggt  | atcttcttca  | ctgcattcta  | acaaataaag  | caatagcatc  | acaaatttca  | 8340 |
| caaataaagc  | ctgtataccg  | tcgacctcta  | gttgtgggtt  | gtccaaactc  | atcaatgtat  | 8400 |
| cttatcatgt  | gtgaaattgt  | tatccgctca  | gctagagctt  | ggcgtaatca  | tggtcatagc  | 8460 |
| tgtttcctgt  | agcctgggg   | gcctaattgag | caattccaca  | caacatacga  | gcccgaagca  | 8520 |
| taaagtgtaa  | tttccagtcg  | ggaaacctgt  | tgagctaact  | cacattaatt  | gcgttgcgct  | 8580 |
| cactgccccg  | agggcggttg  | cgatttgggc  | cgtgccagct  | gcattaatga  | atcgcccaac  | 8640 |
| gcgcggggag  | cgttcggctg  | cggcgagcgg  | gctcttccgc  | ttcctcgtc   | actgactcgt  | 8700 |
| tgcgctcggt  | atcaggggat  | aagcgaggaa  | tatcagctca  | ctcaaaggcg  | gtaatacggg  | 8760 |
| tatccacaga  | taaaaaggcc  | gcgttgctgg  | agaacatgtg  | agcaaaaggc  | cagcaaaagg  | 8820 |
| ccaggaaccg  | aaatcgacgc  | tcgaagtcaga | cgtttttcca  | taggctccgc  | ccccctgacg  | 8880 |
| agcatcacaa  | tccccctgga  | agctccctcg  | tgcgctctcc  | tggtccgacc  | ctataaagat  | 8940 |
| accaggcggt  | gtccgccttt  | ctcccttcgg  | gaagcgtggc  | gctttctcaa  | ctgcccgtta  | 9000 |
| ccggataacct | cagttcgggtg | taggtcggtc  | gctccaaagt  | gggctgtgtg  | tgctcacgct  | 9060 |
| gtaggtatct  | cgaccgctgc  | gccttatccg  | gtaactatcg  | tcttgagtcc  | cacgaacccc  | 9120 |
| ccgttcagcc  | atcgccactg  | gcagcagcca  | ctggtaacag  | gattagcaga  | aaccgggtaa  | 9180 |
| gacacgactt  | tacagagttc  | ttgaagtgg   | ggcctaacta  | cggctacact  | gcgaggtag   | 9240 |
| taggcgggtg  | ctgcgctctg  | ctgaagccag  | ttaccttcgg  | aaaaagagtt  | agaaggacag  | 9300 |
| tatttggtat  | acaaaccacc  | gctggtagcg  | gtggtttttt  | tggttgcaag  | ggtagctctt  | 9360 |
| gatccggcaa  | aaaaggatct  | caagaagatc  | ctttgatctt  | ttctacgggg  | cagcagatta  | 9420 |
| cgcgagaaaa  | aaactcacgt  | taagggattt  | tggtcatgag  | attatcaaaa  | tctgacgctc  | 9480 |
| agtggaaacga | tttaaattaa  | aaatgaagtt  | ttaatcaat   | ctaaagtata  | aggatcttca  | 9540 |
| cctagatcct  | cagttacca   | tgcttaata   | gtgaggcacc  | tatctcagcg  | tatgagtaaa  | 9600 |
| cttgggtctga |             |             |             |             | atctgtctat  | 9660 |

```

ttcgttcata catagtgtgc tgactccccg tcgtgtagat aactacgata cgggaggggt 9720
taccatctgg cccagtgct gcaatgatac cgcgagaccc acgtcaccg gctccagatt 9780
tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 9840
ccgctccat ccagtctatt aattgttgcc ggaagctag agtaagtagt tcgccagtta 9900
atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg 9960
gtatggcttc attcagctcc gggtcccaac gatcaaggcg agttacatga tccccatgt 10020
tgtgcaaaaa agcgggttagc tccttcgggtc ctccgatcgt tgtcagaagt aagtggccg 10080
cagtgttate actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg 10140
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 10200
ggcgaccgag ttgctcttgc ccggcgtaaa tacgggataa taccgcgcca catagcagaa 10260
ctttaaaagt gctcatcatt ggaaaaacgtt cttcggggcg aaaactctca aggatcttac 10320
cgctgttgag atccagttcg atgtaaccca ctctgtgcacc caactgatct tcagcatctt 10380
ttactttcac cagcgtttct ggttgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 10440
gaataagggc gacacggaaa tgttgaatac tcatactctt cttttttcaa tattattgaa 10500
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 10560
aacaaatagc ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 10610

```

<210> 17  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 17  
 tgtacaccgg atccggcgca cacc

24

<210> 18  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 18  
 cacaacgagc tcaattaatt aattgccaca tcttc

35

<210> 19  
 <211> 4  
 <212> PRT  
 <213> adenovirus

<400> 19  
 Thr Leu Trp Thr  
 1

<210> 20  
 <211> 12  
 <212> PRT  
 <213> adenovirus

<400> 20  
 Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser  
 1 5 10

<210> 22  
 <211> 327  
 <212> DNA

## &lt;213&gt; adenovirus

&lt;400&gt; 22

```

agatctgaat tgcagctcgc tgttgggctc gcggttgagg acaaactctt cgcggtcttt 60
ccagtactct tggatcggaa acccgtcggc ctccgaacgg tactccgccca ccgagggacc 120
tgagcgagtc cgcacgcacc ggatcggaaa acctctcgag aaaggcgtct aaccagtcac 180
agtcgcaagg taggctgagc accgtggcgg gcggcagcgg gtggcggtcg gggttgtttc 240
tggcgagggt gctgctgatg atgtaattaa agtaggcggt cttgagacgg cggatggtcg 300
aggtagggtg tggcaggctt gagatct

```

&lt;210&gt; 23

&lt;211&gt; 32480

&lt;212&gt; DNA

&lt;213&gt; adenovirus

&lt;400&gt; 23

```

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60
ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180
gtgtgcgcgg gtgtacacag gaagtacaaa ttttcgcgcg gtttttaggcg gatgttgtag 240
taaattttgg cgtaaccgag taagatttgg ccatttttcg gggaaaactg aataagagg 300
agtgaactct gaataatttt gtgttactca tagcgcgtaa tctctagcat cgatgtcgac 360
aagcttgaat tgcattaatg tgagttagct cactcattag gcaccccagg ctttacactt 420
tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa 480
cagctatgac catgattacg aattcggcgc agcaccatgg cctgaaataa cctctgaaag 540
aggaacttgg ttaggtacct tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt 600
agggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca tgcattctaa 660
ttagtcagca accaggtgtg gaaagtcccc aggcctcccc gcaggcagaa gtatgcaaag 720
catgcatctc aattagtcag caaccatagt cccgcccta actccgccca tccccccct 780
aactccgccc agttccgccc attctccgcc ccattggctg ctaatttttt ttatttatgc 840
agaggccgag gccgcctcgg cctctgagct attccagaag tagtgaggag gcttttttgg 900
aggcctaggc ttttgcaaaa agcttgggat ctctataatc tcgcgcaacc tattttcccc 960
tcgaacactt ttttaagcgg agataaacag gctgggacac ttcacatgag cgaaaaatac 1020
atcgtcacct gggacatgtt gcagatccat gcacgtaaac tcgcaagccg actgatgcct 1080
tctgaacaat ggaaaggcat tattgccgta agccgtggcg gtctgggtacc ggtgggtgaa 1140
gaccagaaac agcacctcga actgagccgc gatattgccc agcgtttcaa cgcgtgtat 1200
ggcgagatcg atcccgtcg tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 1260
cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 1320
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt tgccctggtt 1380
ccggcaccag aagcgggtgc ggaagctgg gcagatgcac gggttacgat cgcccatcta caccacgta 1500
gtcgtcgtcc cctcaaaactg gcagatgcac ggttcccacg agaatccgac ggggttgttac 1560
acctatccca ttacggtcaa tccgcgcttt ctacaggaag gccagacgcg aattattttt 1620
tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gctgggtcgg ttacggccag 1680
gatggcgtaa actcggcggt tcatctgtgg tgcaacgggc gctgggtcgg ttacggccag 1740
gacagtcgtt tgccgtctga atttgacctg agcgcatttt tacgcgccgg agaaaaccgc 1800
ctcgcggtga tgggtgctgag ttggagtgcg ggcagttatc tggaagatca ggatagtgtg 1860
cggatgagcg gcattttccg tgacgtctcg ttgctgcata aaccgactac acaaatcagc 1920
gatttccatg ttgccaactc ctttaatgat gatttcagcc ggcgtgtact ggaggctgaa 1980
gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa cagtttcttt atggcagggt 2040
gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg aaattatcga tgagcgtgg 2100
ggttatgccc atcggtcac actacgtctg aacgtcgaaa acccgaaact gtggagcgcc 2160
gaaatcccga atctctatcg tgcggtggtt gaactgcaca ccgccgacgg cacgctgatt 2220
gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga ttgaaaatgg tctgtgtctg 2280
ctgaacggca agcgtttgct gattcgaggg gttaaccgtc acgagcatca tctctgcat 2340
ggtcaggta tggtatgagc gacgatgggt gcattatccg tgctgatgaa gcagaacaac 2400
tttaacggcg tgcgctgttc gcatatccg aaccatccgc tgtgttacac gctgtgcgac 2460
cgctacggcc tgtatgtggt ggatgaagcc aatatgaaa cccacggcat ggtgccaatg 2520
aatcgtctga ccgatgatcc gcgctggcta ccggcgatga gcgaacgcgt aacgcgaatg 2580
gtgcagcgcg atcgtaatca cccgagtggt atcatctggt cgctggggaa tgaatcaggc 2640
cacggcgcta atcacgacgc gctgtatcgc tggatcaaat ctgtcgatcc tccccgccg 2700
gtgcagtatg aaggcgcgcg agccgacacc acggccaccg atattatttg cccgatgtac 2760
gcgcgcgtgg atgaagacca gcccttcccc gctgtgcgga aatggtccat caaaaaatgg

```

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ctttcgctac  | ctggagagac  | gcgcccgtcg  | atcctttgcg  | aatacgccca  | cgcgatgggt  | 2820 |
| aacagtcttg  | gcggtttcgc  | ttaaatactgg | caggcggtttc | gtcagtatcc  | ccggtttacag | 2880 |
| ggcggcttcg  | tctgggactg  | ggtggatcag  | tcgctgatta  | aatatgatga  | aaacggcaac  | 2940 |
| ccgtggctcg  | cttacggcgg  | tgattttggc  | gatacgccga  | acgatcgcca  | gttctgtatg  | 3000 |
| aacggctcgg  | tctttggcga  | ccgcacgcgc  | catccagcgc  | tgacggaagc  | aaaacaccag  | 3060 |
| cagcagtttt  | tccagttccg  | tttatccggg  | caaaccatcg  | aagtgaccag  | cgaataacctg | 3120 |
| ttccgctata  | gcgataaacga | gctcctgcac  | tggtatgggtg | cgctggatgg  | taagccgctg  | 3180 |
| gcaagcgggtg | aagtgcctct  | ggatgtcgct  | ccacaaggta  | aacagttgat  | tgaactgcct  | 3240 |
| gaactaccgc  | agccggagag  | cgccggggcaa | ctctgggtca  | cagtacgcgt  | agtgaaccgc  | 3300 |
| aacgcgaccg  | catggtcaga  | agccggggcac | atcagcgccct | ggcagcagtg  | gcgtctggcg  | 3360 |
| gaaaacctca  | gtgtgacgct  | ccccgcgcgc  | tcccacgcca  | tcccgcactct | gaccaccagc  | 3420 |
| gaaatggatt  | tttgcatacga | gctgggtaaat | aagcgttggc  | aatttaaccg  | ccagtcaggc  | 3480 |
| tttctttcac  | agatgtggat  | tggtcgataaa | aaacaactgc  | tgacgccgct  | gcgcgatcag  | 3540 |
| ttcaccgctg  | caccgctgga  | taacgacatt  | ggcgtaagtg  | aagcgaccgc  | cattgacctg  | 3600 |
| aacgcctggg  | tcgaacgctg  | gaaggcggcg  | ggccattacc  | aggccgaagc  | agcgttggtg  | 3660 |
| cagtgcacgg  | cagatacact  | tgctgatcgc  | gtgctgatta  | cgaccgctca  | cgcggtggcag | 3720 |
| catcagggga  | aaaccttatt  | tatcagccgg  | aaaacctacc  | ggattgatgg  | tagtggtcaa  | 3780 |
| atggcgatta  | cggttgatgt  | tgaagtggcg  | agcgatacac  | cgcatccggc  | gcggtattggc | 3840 |
| ctgaactgcc  | agctggcgca  | ggtagcagag  | cggtgaaact  | ggctcggatt  | agggccgcaa  | 3900 |
| gaaaactatc  | ccgaccgcct  | tactgccgcc  | tgttttgacc  | gctgggatct  | gccattgtca  | 3960 |
| gacatgtata  | ccccgtacgt  | cttcccgcgc  | gaaaacggtc  | tgcgctcgcg  | gacgcgcgaa  | 4020 |
| ttgaattatg  | gcccacacca  | gtggcgcgcc  | gacttccagt  | tcaacatcag  | ccgctacagt  | 4080 |
| caacagcaac  | tgatggaaac  | cagccatcgc  | catctgctgc  | acgcggaaga  | aggcacatgg  | 4140 |
| ctgaatatcg  | acggtttcca  | tatggggatt  | ggtggcgacg  | actcctggag  | cccgatgta   | 4200 |
| tcggcggaat  | tccagctgag  | cgccgggtcg  | taccattacc  | agttaggtctg | gtgtcaaaaa  | 4260 |
| taataataac  | cgggcaggcc  | atgtctgccc  | gtatttcgcg  | taaggaaatc  | cattatgtac  | 4320 |
| tatttaaaaa  | acacaaaactt | ttggatgttc  | ggttttattct | ttttctttta  | cttttttatc  | 4380 |
| atgggagcct  | acttcccgtt  | tttcccgaat  | ttggtacatg  | acatcaacca  | tatcagcaaa  | 4440 |
| agtgatacgg  | gtattatttt  | tgccgctatt  | tctctgttct  | cgctattatt  | ccaaccgctg  | 4500 |
| tttggctctgc | tttctgacaa  | actcggaact  | tgttttattgc | agcttataat  | ggttacaaat  | 4560 |
| aaagcaatag  | catcacaaat  | ttcacaaata  | aagcattttt  | ttcactgcat  | tctagtttgt  | 4620 |
| gtttgtccaa  | actcatcaat  | gtatcttatc  | atgtctggat  | ccagatctgg  | gcgtggctta  | 4680 |
| aggggtgggaa | agaatatata  | aggtgggggt  | cttatgtagt  | tttgtatctg  | ttttgcagca  | 4740 |
| gccgcgcgcg  | ccatgagcac  | caactcgttt  | gatggaagca  | ttgtgagctc  | atatttgaca  | 4800 |
| acgcgcactgc | ccccatgggc  | cggggtgcgt  | cagaatgtga  | tgggctccag  | catttgatgg  | 4860 |
| cgccccgtcc  | tgccccgaaa  | ctctactacc  | ttgacctacg  | agaccgtgtc  | tggaaacgcg  | 4920 |
| ttggagactg  | cagcctccgc  | cgccgcttca  | gccgctgcag  | ccaccgcccgc | cggtattgtg  | 4980 |
| actgactttg  | ctttcctgag  | cccgtttgca  | agcagtgcag  | cttcccgttc  | atccgcccgc  | 5040 |
| gatgacaagt  | tgacggctct  | tttggcacaa  | ttggattctt  | tgaccgggga  | acttaattgc  | 5100 |
| gtttctcagc  | agctgttggg  | tctgcgccag  | caggtttctg  | ccctgaaggc  | ttcctccctc  | 5160 |
| cccaatgcgg  | tttaaaacat  | aaataaaaaa  | ccagactctg  | tttggatttg  | gatcaagcaa  | 5220 |
| gtgtccttgct | gtctttattt  | aggggttttg  | cgcgcgcggg  | aggcccgggg  | ccagcgggtc  | 5280 |
| cggtcgttga  | gggtcctgtg  | tattttttcc  | aggacgtggg  | aaagggtgact | ctggatgttc  | 5340 |
| agatacatgg  | gcataagccc  | gtctctgggg  | tggaggtagc  | accactgcag  | agcttcatgc  | 5400 |
| tgccgggttg  | tggtttagat  | gatccagtcg  | tagcaggagc  | gctgggcgtg  | gtgcctaaaa  | 5460 |
| atgtctttca  | gtagcaagct  | gattgccagg  | ggcaggccct  | tggtgttaagt | gtttacaaag  | 5520 |
| cggttaagct  | gggatgggtg  | catacgtggg  | gatatgagat  | gcactcttga  | ctgtattttt  | 5580 |
| aggttggcta  | tggtcccgag  | catatccctc  | cggggattca  | tgttgtgcag  | aaccaccagc  | 5640 |
| acagtgtatc  | cggtgcactt  | gggaaatttg  | tcatgtagct  | tagaaggaaa  | tgcggtggaag | 5700 |
| aacttgga    | cgcccttggtg | acctccaaga  | ttttccatgc  | attcgtccat  | aatgatggca  | 5760 |
| atgggcccac  | gggcggcggc  | ctgggcgaag  | atattttctg  | gatcactaac  | gtcatagttg  | 5820 |
| tggtccagga  | tgagatcgtc  | ataggccatt  | tttacaaagc  | gcgggcggag  | gggtgccagac | 5880 |
| tgccgtataa  | tggttccatc  | cggcccaggg  | gcgtagttac  | cctcacagat  | ttgcatttcc  | 5940 |
| cacgctttga  | gttcagatgg  | ggggatcatg  | tctacctgcg  | gggcgatgaa  | gaaaacgggt  | 6000 |
| tccggggtag  | gggagatcag  | ctgggaagaa  | agcaggttcc  | tgagcagctg  | cgacttaccg  | 6060 |
| cagccgggtg  | gcccgtaaat  | cacacctatt  | accgggtgca  | actggtagtt  | aagagagctg  | 6120 |
| cagctgcctg  | catccctgag  | cagggggggc  | acttcgttaa  | gcattgtcct  | gactcgcatg  | 6180 |
| ttttccctga  | ccaaatccgc  | cagaaggcgc  | tcgccgccca  | gcgatagcag  | ttcttgcaag  | 6240 |
| gaagcaaagt  | ttttcaacgg  | tttgagaccg  | tccgcctgag  | gcattgcttt  | gagcgtttga  | 6300 |
| ccaagcagtt  | ccaggcggtc  | ccacagctcg  | gtcacctgct  | ctacggcatc  | togatccagc  | 6360 |
| atatctcttc  | gtttcgcggg  | ttggggcggc  | tttcgctgta  | cggcagtagt  | cggtgctcgt  | 6420 |
| ccagacgggc  | cagggtcatg  | tctttccacg  | ggcgcagggt  | cctcgtcagc  | gtagtctggg  | 6480 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tcacgggtgaa | gggggtgcgct | ccggggtgcg  | cgctggccag  | ggtgcgcttg  | aggctgggtcc | 6540  |
| tgctgggtgct | gaagcgctgc  | cggtcttcgc  | cctgcgcgct  | ggccaggtag  | catttgacca  | 6600  |
| tggtgtcata  | gtccagcccc  | tccgcggcgt  | ggcccttggc  | gcgcagcttg  | cccttggagg  | 6660  |
| aggcgccgca  | cgagggggcag | tgacagacttt | tgagggcgta  | gagcttgggc  | gcgagaaata  | 6720  |
| ccgattccgg  | ggagtaggca  | tccgcgcgcg  | aggccccgca  | gacgggtctcg | cattccacga  | 6780  |
| gccaggtgag  | ctctggccgt  | tccgggtcaa  | aaaccagggt  | tcccccatgc  | tttttgatgc  | 6840  |
| gtttcttacc  | tctgggttcc  | atgagccggt  | gtccacgctc  | ggtgacgaaa  | aggctgtccg  | 6900  |
| tgctcccgta  | tacagacttg  | agaggcctgt  | cctcgagcgg  | tggtccgcgg  | tcctcctcgt  | 6960  |
| atagaaactc  | ggaccactct  | gagacaaaagg | ctcgcgctcca | ggccagcacg  | aaggaggcta  | 7020  |
| agtgggaggg  | gtagcgggtcg | ttgtccacta  | gggggtccac  | tcgctccagg  | gtgtgaagac  | 7080  |
| acatgtcgcc  | ctcttcggca  | tcaaggaaagg | tgattgggtt  | gtaggtgtag  | gccacgtgac  | 7140  |
| cggtgttcc   | tgaagggggg  | ctataaaaagg | gggtgggggc  | gcgttcgtcc  | tcactctctt  | 7200  |
| ccgcatcgct  | gtctgcgagg  | gccagctgtt  | gggggtgagta | ctccctctga  | aaagcggggca | 7260  |
| tgacttctgc  | gctaagattg  | tcagtttcca  | aaaacgagga  | ggatttgata  | ttcacctggc  | 7320  |
| ccgcggtgat  | gcctttgagg  | gtggccgcat  | ccatctggtc  | agaaaagaca  | atctttttgt  | 7380  |
| tgtcaagctt  | ggtggcaaac  | gacccgtaga  | gggctgtgga  | cagcaacttg  | gcgatggagc  | 7440  |
| gcagggtttg  | gtttttgtcg  | cgatcggcgc  | gctccttggc  | cgcgatgttt  | agctgcacgt  | 7500  |
| attcgcgcgc  | aacgcaccgc  | cattcgggaa  | agacgggtgt  | gcgctcgctc  | ggcaccaggc  | 7560  |
| gcacgcgcca  | accgcggttg  | tgacgggtga  | caaggccaac  | gctggtggct  | acctctccgc  | 7620  |
| gtaggcgctc  | gttgggtccag | cagaggcgcc  | cgcccttgcg  | cgagcagaat  | ggcggtaggg  | 7680  |
| ggtctagctg  | cgtctcgctc  | gggggggtctg | cgtccacggt  | aaagaccccg  | ggcagcaggc  | 7740  |
| gcgcgtcgaa  | gtagtctatc  | ttgcacacct  | gcaagcttag  | cgctcgctgc  | catgcgcggg  | 7800  |
| cgggaagcgc  | gcgctcgat   | gggttgagtg  | ggggacccca  | tgcatggggg  | tgggtgagcg  | 7860  |
| cggaggcgta  | catgccgcaa  | atgtcgtaaa  | cgtagagggg  | ctctctgagt  | attccaagat  | 7920  |
| atgtagggtg  | gcatcttcca  | ccgcggatgc  | tggcgcgcac  | gtaatcgat   | agttcgtcgc  | 7980  |
| agggcgtag   | gaggtcggga  | ccgaggttgc  | tacggggcgg  | ctgctctgct  | cggaagacta  | 8040  |
| tctgcctgaa  | gatggcatgt  | gagttggatg  | atatggttgg  | acgctggaag  | acgttgaagc  | 8100  |
| tggcgctctgt | gagacctacc  | gcgtcacgca  | cgaaggaggc  | gtaggagtgc  | cgcagcttgt  | 8160  |
| tgaccagctc  | ggcggtgacc  | tgacagctta  | gggcgcagta  | gtccagggtt  | tccttgatga  | 8220  |
| tgataactt   | atcctgtccc  | ttttttttcc  | acagctcgcg  | gttgaggaca  | aactcttcgc  | 8280  |
| ggtctttcca  | gtactcttgg  | atcggaaaacc | cgctcgccctc | cgaacggtaa  | gagcctagca  | 8340  |
| tgtagaactg  | gttgacggcc  | tggtaggcgc  | agcatccctt  | ttctacgggt  | agcgcgtatg  | 8400  |
| cctgcgcggc  | cttccggagc  | gaggtgtggg  | tgagcgcaaa  | ggtgtccctg  | accatgactt  | 8460  |
| tgaggtagctg | gtatttgaa   | tcagtgtcgt  | cgcatccgcc  | ctgctcccag  | agcaaaaagt  | 8520  |
| ccgtgcgctt  | tttggaaacgc | ggatttggca  | gggcgaagg   | gacatcggtg  | aagagtatct  | 8580  |
| ttcccgcgcg  | aggcataaag  | ttgcgtgtga  | tgcggaagg   | tcccgccacc  | tcggaaaggc  | 8640  |
| tgtaatttac  | ctgggcggcg  | agcacgatct  | cgctcaaagcc | gttgatgttg  | tggcccacaa  | 8700  |
| tgtaaagttc  | caagaagcgc  | gggatgccct  | tgatggaagg  | caatttttta  | agttcctcgt  | 8760  |
| aggtgagctc  | ttcagggggag | ctgagcccg   | gctctgaaag  | ggcccagctc  | gcaagatgag  | 8820  |
| ggttggaaag  | gacgaatgag  | ctccacaggt  | cacgggccat  | tagcatttgc  | aggtggctgc  | 8880  |
| gaaaggtcct  | aaactggcga  | cctatggcca  | ttttttctgg  | ggtgatgcag  | tagaaggtaa  | 8940  |
| gcgggtcttg  | ttcccagcgg  | tcccatccaa  | ggttcgcggc  | taggtctcgc  | gcggcagtca  | 9000  |
| ctagaggctc  | atctccgcgg  | aacttcatga  | ccagcatgaa  | gggcacgagc  | tgcttcccaa  | 9060  |
| aggcccccat  | ccaagtatag  | gtctctacat  | cgtaggtgac  | aaagagacgc  | tcggtgcgag  | 9120  |
| gatgcgagcc  | gatcggaag   | aactggatct  | cccgccacca  | attggaggag  | tggctattga  | 9180  |
| tgtggtgaaa  | gtagaagtcc  | ctgcgacggg  | ccgaacactc  | gtgctggctt  | ttgtaaaaac  | 9240  |
| gtgcgcagta  | ctggcagcgg  | tgacgggtg   | gtacatcctg  | cacgaggttg  | acctgacgac  | 9300  |
| cgcgacacaag | gaagcagagt  | gggaatttga  | gccccctgcc  | tggcgggttt  | ggctggtggt  | 9360  |
| cttctacttc  | ggctgcttgt  | ccttgaccgt  | ctggctgctc  | gaggggagtt  | acggtggatc  | 9420  |
| ggaccaccac  | gccgcgcgag  | cccaaagtcc  | agatgtccgc  | gcgcggcggt  | cgagccttga  | 9480  |
| tgacaacatc  | ctgcaggttt  | acctcgcata  | tggtctggag  | ctcccgcggc  | gtcaggtcag  | 9540  |
| gcgggagctc  | gtgcaggttt  | acctcgcata  | gacgggtcag  | ggcgcgggct  | agatccaggt  | 9600  |
| gataccta    | ttccaggggc  | tggttgggtg  | cggtgtcgat  | ggcttgcaag  | aggcgcgcat  | 9660  |
| cccgcgccgc  | gactacggta  | ccgcgcggcg  | ggcggtgggc  | cgcggggggtg | tccttggatg  | 9720  |
| atgcacttaa  | aagcgggtgac | gcgggcgagc  | ccccggaggt  | aggggggggt  | ccggaccgcg  | 9780  |
| cgggagaggg  | ggcaggggca  | cgtcggcgcc  | gcgcgcgggc  | aggagctggt  | gctgcgcgcg  | 9840  |
| taggttgctg  | gcgaacgcga  | cgacgcggcg  | ggtgatctcc  | tgaatctggc  | gcctctgcgt  | 9900  |
| gaagacgacg  | ggcccgggtga | gcttgagcct  | gaaagagagt  | tcgacagaat  | caatttcggt  | 9960  |
| gtcgttgacg  | gcggcctggc  | gcacacgtct  | ctgcacgtct  | cctgagttgt  | cttgataggc  | 10020 |
| gatctcggcc  | atgaactgct  | cgatctcttc  | ctcctggaga  | tctccgcgctc | cggctcgctc  | 10080 |
| cacgggtggcg | gcgaggtcgt  | tggaaatgcg  | ggccatgagc  | tgcgagaagg  | cgttgaggcc  | 10140 |
| tccctcgcttc | cagacgcggc  | tgtagaccac  | gcccccttcg  | gcacgcggg   | cgcgcatgac  | 10200 |

|             |            |             |             |            |             |       |
|-------------|------------|-------------|-------------|------------|-------------|-------|
| cacctgcgcg  | agattgagct | ccacgtgccc  | ggcgaagacg  | gcgtagtttc | gcaggcgctg  | 10260 |
| aaagaggtag  | ttgaggggtg | tggcgggtgt  | ttctgcccac  | aagaagtaca | taacccagcg  | 10320 |
| tcgcaacgtg  | gatttcgttg | tatcccccaa  | ggcctcaagg  | cgctccatgg | cctcgtagaa  | 10380 |
| gtccacggcg  | aagttgaaaa | actgggagtt  | gcgcgccgac  | acggttaact | cctcctccag  | 10440 |
| aagacggatg  | agctcggcga | cagtgtcgcg  | cacctcgcg   | tcaaaggcta | caggggcctc  | 10500 |
| ttcttcttct  | tcaatctcct | cttcataaag  | ggcctcccc   | tcttcttctt | ctggcgccgg  | 10560 |
| tgggggaggg  | gggacacggc | ggcgacgacg  | ggcgaccggg  | aggcggtcga | caaagcgctc  | 10620 |
| gatcatctcc  | ccgcggcgac | ggcgcatggg  | ctcggtgacg  | gcgcggccgt | tctcgcgggg  | 10680 |
| gcgcagttgg  | aagacgccc  | ccgtcatgtc  | ccggttatgg  | gttggcgggg | ggctgccatg  | 10740 |
| cggcagggat  | acggcgctaa | cgatgcacat  | caacaattgt  | tgtgtaggta | ctccgcggcc  | 10800 |
| gagggacctg  | agcgagtcgg | catcgaccgg  | atcggaatac  | ctctcgagaa | aggcgctctaa | 10860 |
| ccagtcacag  | tcgcaaggta | ggctgagcac  | cgtagcgggc  | ggcagcgggc | ggcggtcggg  | 10920 |
| gttgtttctg  | gcgaggggtg | tgctgatgat  | gtaattaaag  | taggcgggtc | tgagacggcg  | 10980 |
| gatggtcgac  | agaagcacca | tgctcctggg  | tccggcctgc  | tgaatgcgca | ggcggtcggc  | 11040 |
| catgccccag  | gcttcgtttt | gacatcgggc  | caggtccttg  | tagtagtctt | gcatgagcct  | 11100 |
| ttctaccggc  | acttcttctt | ctccttcttc  | ttgtcttgca  | tctcttgcat | ctatcgctgc  | 11160 |
| ggcgcgggcg  | gagtttggtc | gtaggtggcg  | ccctcttctt  | cccatgctgt | tgaccccgaa  | 11220 |
| gcccctcatc  | gggttaggtc | gggttaggtc  | ggcgacaacg  | cgctcggtta | atatggcctg  | 11280 |
| ctgcacctgc  | gtgagggtag | actggaagtc  | atccatgtcc  | acaaagcggt | ggtatgcgcc  | 11340 |
| cgtgttgatg  | gtgtaagtgc | agttggccat  | aacggaccag  | ttaacgggtc | ggtgacccgg  | 11400 |
| ctgcgagagc  | tcggtgtacc | tgagacgcga  | gtaagccctc  | gagtcaaata | cgtagtcgtt  | 11460 |
| gcaagtccgc  | accaggtact | ggtatcccac  | caaaaagtgc  | ggcgggcggt | ggcggtagag  | 11520 |
| gggccagcgt  | agggtggccg | gggctccggg  | ggcgagatct  | tccaacataa | ggcgatgata  | 11580 |
| tccgtagatg  | tacctggaca | tccaggtgat  | ggcgggcgcg  | gtggtggagg | cgcgcggaag  | 11640 |
| gtcgcggacg  | cggttccaga | tggtgcgcag  | cggcaaaaag  | tgctccatgg | tcgggacgct  | 11700 |
| ctggccgggtc | aggcgcgcg  | aatcgttgac  | gctctagacc  | gtgcaaaaag | agagcctgta  | 11760 |
| agcgggcact  | cttccgtggt | ctgggtggata | aatctcgcaag | ggtatcatgg | cggacgaccg  | 11820 |
| gggttcgagc  | cccgtatccg | gcccgtccgc  | gtatccatg   | cggttaccgc | cccggtgtcg  | 11880 |
| aaccacaggt  | tcgcacgtca | gacaacgggg  | gagtgctcct  | tttggtcttc | ttccaggcgc  | 11940 |
| ggcggtgct   | gcgctagctt | ttttggccac  | tgccgcgcgc  | cagcgtaagc | ggttaggctg  | 12000 |
| gaaagcgaaa  | gcattaagtg | gctcgctccc  | tgtagccgga  | gggttatttt | ccaagggttg  | 12060 |
| agtccgggga  | ccccgggttc | gagttctcga  | cggcccgagc  | tgccggcgaa | gggggtttgc  | 12120 |
| ctcccgtca   | tgcaagaccc | cgcttgcaaa  | ttcctccgga  | aacagggacg | agcccccttt  | 12180 |
| ttgcttttcc  | cagatgcac  | cggtgtcg    | gcagatgcgc  | ccccctcctc | agcagcggca  | 12240 |
| agagcaagag  | cagcggcaga | catgcagggc  | accctcccc   | cctcctaccg | cgtcaggagg  | 12300 |
| ggcgacatcc  | cgggttgacg | cggcagcaga  | tggtgattac  | gaacccccgc | ggcgccgggg  | 12360 |
| ccggcactac  | ctggacttgg | aggagggcga  | gggcttggg   | cggctaggag | cgccctctcc  | 12420 |
| tgagcggtag  | ccaagggtgc | agctgaagcg  | tgatacgcgt  | gaggcgtag  | tgcccgcgga  | 12480 |
| gaacctgttt  | cgcgaccgcg | aggagagagga | gcccagggag  | atgcgggatc | gaaagtcca   | 12540 |
| cgcagggcgc  | gagctgcggc | atggcctgaa  | tcgcgagcgg  | ttgctgcgcg | aggaggactt  | 12600 |
| tgagcccagc  | gcgcgaaacc | ggattagtc   | cgcgcgcgca  | cacgtggcgg | ccgcccacct  | 12660 |
| ggtaacgcga  | tacgagcaga | cggtgaacca  | ggagattaac  | tttcaaaaaa | gctttaacaa  | 12720 |
| ccacgtgcgt  | acgcttggtg | cgcgcgagga  | ggtggctata  | ggactgatgc | atctgtggga  | 12780 |
| ctttgtgaag  | gcgctggagc | aaaacccaaa  | tagcaagccg  | ctcatggcgc | agctgttctt  | 12840 |
| tatagtgcag  | cacagcaggg | acaacgaggg  | attcagggat  | gcgctgctaa | acatagtaga  | 12900 |
| gcccaggggc  | cgctggctgc | tcgatttgat  | aaacatcctg  | cagagcatag | tggtgcagga  | 12960 |
| gcgagcttg   | agcctggctg | acaagggtgg  | cgccatcaac  | tattccatgc | ttagcctggg  | 13020 |
| caagttttac  | gcccgcgaag | tataccatac  | cccttacgtt  | cccatagaca | aggaggtaaa  | 13080 |
| gatcgagggg  | ttctacatgc | gcattggcgt  | gaaggtgctt  | accttgagcg | acgacctggg  | 13140 |
| cgtttatcgc  | aacgagcgca | tcacacaagg  | cgtgagcgtg  | agccggcggg | gcgagctcag  | 13200 |
| cgaccgcgag  | ctgatgcaca | gcctgcaaa   | ggccttggtt  | ggcacgggca | gcggcgatag  | 13260 |
| agaggccgag  | tcctactttg | acgcggggcg  | tgacctgcgc  | tgggccccaa | gccgacgcgc  | 13320 |
| cctggaggga  | gctggggccg | gacctgggct  | ggcggtggca  | cccgcgcgcg | ctggcaacgt  | 13380 |
| cggcgggcgtg | gaggaatatg | acgaggacga  | tgagtacgag  | ccagaggacg | gcgagtacta  | 13440 |
| agcgggtgatg | tttctgatca | gatgatgcaa  | gacgcaacgg  | acccggcggt | gcggggcgcg  | 13500 |
| ctgcagagcc  | agccgtccgg | ccttaactcc  | acggacgact  | ggcgccaggt | catggaccgc  | 13560 |
| atcatgtcgc  | tgactgcgcg | caatcctgac  | gcgttccggc  | agcagccgca | ggccaaccgg  | 13620 |
| ctctccgcaa  | ttctggaagc | ggtgggtccc  | gcgcgcgcaa  | acccacgcga | cgagaagggtg | 13680 |
| ctggcgatcg  | taaacgcgct | ggccgaaaac  | agggccatcc  | ggcccagcga | ggccggcctg  | 13740 |
| gtctacgacg  | cgctgcttca | gcgcgtggct  | cgttacaaca  | gcggcaacgt | gcagaccaac  | 13800 |
| ctggacccgc  | tgggggggga | tgtgcgcgag  | ccgtggcgcg  | agcgtgagcg | cgcgccagag  | 13860 |
| caggggcaacc | tgggctccat | ggttgactaa  | aacgccttcc  | tgagtacaca | gcccgcgaac  | 13920 |

|             |             |             |            |            |             |       |
|-------------|-------------|-------------|------------|------------|-------------|-------|
| gtgccgcggg  | gacaggagga  | ctacaccaac  | tttgtgagcg | caactgcggt | aatggtgact  | 13980 |
| gagacaccgc  | aaagtggagt  | gtaccagctc  | gggccagact | atgttttcca | gaccagtaga  | 14040 |
| caaggcctgc  | agaccgtaaa  | cctgagccag  | gctttcaaaa | acttgacagg | gctgtggggg  | 14100 |
| gtgcgggctc  | ccacaggcga  | ccgcgcgacc  | gtgtctagct | tgctgacgcc | caactcgcgc  | 14160 |
| ctgttgctgc  | tgctaatagc  | gcccttcacg  | gacagtggca | gcgtgtcccg | ggacacatac  | 14220 |
| ctaggtcact  | tgctgacact  | gtaccgcgag  | gccatagggt | aggcgcatgt | ggacgagcat  | 14280 |
| actttccagg  | agattacaag  | tgctagccgc  | gcgtgggggc | aggaggacac | gggcagcctg  | 14340 |
| gaggcaaccc  | taaactacct  | gctgaccaac  | cggcgcgaga | agatccccct | gttgacacagt | 14400 |
| ttaaacagcg  | aggaggagcg  | cattttgccg  | tacgtgcagc | agagcgtgag | ccttaacctg  | 14460 |
| atgcgcgacg  | gggtaacgcc  | cagcgtggcg  | ctggacatga | ccgcgcgcaa | catgggaacc  | 14520 |
| ggcatgtatg  | cctcaaacgc  | gccgtttatc  | aaccgcctaa | tggactactt | gcatcgcgcg  | 14580 |
| gccgcgctga  | accccgagta  | tttcaccaat  | gccatcttga | acccgcactg | gctaccgccc  | 14640 |
| ccttggtttct | acaccggggg  | attcgagggt  | cccgagggtg | acgatggatt | cctctgggac  | 14700 |
| gacatagacg  | acagcgtggt  | ttccccgcaa  | ccgcagaccc | tgctagagtt | gcaacacgcg  | 14760 |
| gagcaggcag  | aggcggcgct  | gcgaaaggaa  | agcttccgca | ggccaagcag | cttgtccgat  | 14820 |
| ctaggcgctg  | cggcccccgc  | gtcagatgct  | agtagcccat | ttccaagctt | gatagggtct  | 14880 |
| cttaccagca  | ctcgcaccac  | ccgcccgcgc  | ctgctgggag | aggaggagta | cctaaacaac  | 14940 |
| tcgctgctgc  | agccgcagcg  | cgaaaaaaac  | ctgcctccgg | catttcccaa | caacgggata  | 15000 |
| gagagcctag  | tggacaagat  | gagtagatgg  | aagacgtacg | cgcaggagca | cagggacgtg  | 15060 |
| ccaggccccg  | gcccccccac  | ccgtcgtcaa  | aggcacgacc | gtcagcgggg | tctggtgtgg  | 15120 |
| gaggacgatg  | actcggcaga  | cgacagcagc  | gtcctggatt | tgggaggggg | tgggaggggg  | 15180 |
| tttgcgccac  | ttcgccccag  | gctggggaga  | atgttttaaa | aaaaaaaaag | catgatgcaa  | 15240 |
| aataaaaaac  | tcaccaaggc  | catggcaccg  | agcgttggtt | ttcttgtatt | ccccttagta  | 15300 |
| tgccggcgcg  | ggcgatgtat  | gaggaagggt  | ctcctccctc | ctacgagagt | gtggtgagcg  | 15360 |
| ggcgcccgcg  | ggcgcgcgcg  | ctgggttctc  | ctcctcgatg | tccccgggac | tcggcgtttg  | 15420 |
| tgccctccgc  | gtacctgccc  | cctaccgggg  | ggagaaacag | catccgttac | tctgagttgg  | 15480 |
| cacccttatt  | cgacaccacc  | cgtgtgtacc  | tgggtggaaa | caagtcaacg | gatgtggcat  | 15540 |
| ccctgaacta  | ccagaacgac  | cacagcaact  | ttctgaccac | gggtcattca | aacaatgact  | 15600 |
| acagcccggg  | gagggcaagc  | acacagacca  | tcaatcttga | cgaccgggtc | cactggggcg  | 15660 |
| gcgacctgaa  | aaccatcctg  | cataccaaca  | tgccaaatgt | gaacgagttc | atgtttacca  | 15720 |
| ataagttaa   | ggcgcggggt  | atggtgtcgc  | gcttgcttac | taaggacaat | caggtggagc  | 15780 |
| tgaatacga   | gtgggtggag  | ttcacgctgc  | ccgagggcaa | ctactccgag | accatgacca  | 15840 |
| tagaccttat  | gaacaacgcg  | atcgtggagc  | actacttgaa | agtgggcaga | cagaacgggg  | 15900 |
| ttctggaaag  | cgacatcggg  | gtaaagtgtg  | acaccgcgaa | cttcagactg | gggtttgacc  | 15960 |
| ccgtcactgg  | tcttgtcatg  | cctgggggtat | atacaaacga | agccttccat | ccagacatca  | 16020 |
| tttgtctgcc  | aggatgcggg  | gtggacttca  | cccacagccg | cctgagcaac | ttgttgggca  | 16080 |
| tccgcaagcg  | gcaacccttc  | caggagggtc  | ttaggatcac | ctacgatgat | ctggagggtg  | 16140 |
| gtaacattcc  | cgcaactgtt  | gatgtggacg  | cctaccaggc | gagcttgaaa | gatgacaccg  | 16200 |
| aacaggggcg  | gggtggcgca  | ggcgcgagca  | acagcagtgg | cagcggcgcg | gaagagaact  | 16260 |
| ccaacgcggc  | agccgcggca  | atgcagccgg  | tggaggacat | gaacgatcat | gccattcgcg  | 16320 |
| gcgacacctt  | tgccacacgg  | gctgaggaga  | agcgcgctga | ggccgaagca | gcggccgaag  | 16380 |
| ctgcgcgccc  | cgctgcgcaa  | cccaggtcgc  | agaagcctca | gaagaaaccg | gtgatcaaac  | 16440 |
| ccctgacaga  | ggacagcaag  | aaacgcagtt  | acaacctaat | aagcaatgac | agcaccttca  | 16500 |
| cccagtaccg  | cagctggtac  | cttgcataca  | actacggcga | ccctcagacc | ggaatccgct  | 16560 |
| catggaccct  | gctttgcact  | cctgacgtaa  | cctgcggctc | ggagcaggtc | tactggtcgt  | 16620 |
| tgccagacat  | gatgcaagac  | cccgtgacct  | ctgcctccac | gcgcagatc  | agcaactttc  | 16680 |
| cggtggtggg  | cgccgagctg  | ttgcccgtgc  | actccaagag | cttctacaac | gaccaggccg  | 16740 |
| tctactccca  | actcatccgc  | cagtttacct  | ctctgaccga | cgtgttcaat | cgctttcccg  | 16800 |
| agaaccagat  | tttgccgcgc  | ccgccagccc  | ccaccatcac | caccgtcagt | gaaaacgttc  | 16860 |
| ctgctctcac  | agatcacggg  | acgctaccgc  | tgcgcaacag | catcgaggga | gtccagcgag  | 16920 |
| tgaccattac  | tgacgcccaga | cgcgcgacct  | gcccctacgt | ttacaaggcc | ctgggcatag  | 16980 |
| tctcgccgcg  | cgctcctatc  | agccgcactt  | tttgagcaag | catgtccatc | cttatatcgc  | 17040 |
| ccagcaataa  | cacaggctgg  | ggcctgcgct  | tcccagcaa  | gatgtttggc | ggggccaaga  | 17100 |
| agcgctccga  | ccaacaccca  | gtgcgcgtgc  | gcgggcaact | ccgcgcgccc | tggggcgcg   | 17160 |
| acaaacgcgg  | ccgcaactgg  | cgcaccaccg  | tcgatgacgc | catcgacgcg | gtggtggagg  | 17220 |
| aggcgcgcaa  | ctacacgccc  | acgcgcgccc  | cagtgctcac | agtggacgcg | gccattcaga  | 17280 |
| ccgtggtgcg  | cggagcccgc  | cgtatgcta   | aaatgaagag | acggcgaggg | cgcgtagcac  | 17340 |
| gtcgccaccg  | ccgcgcgacc  | ggcactgccc  | cccaacgcgc | ggcgggggcc | ctgcttaacc  | 17400 |
| gcgcacgtcg  | caccggccga  | cggcgggcca  | tgccggccgc | tcgaaggctg | gccgcgggta  | 17460 |
| ttgtcactgt  | gccccccagg  | tccaggcgac  | gagcgccgcg | cgcagcagcc | gcggccatta  | 17520 |
| tgctcatgac  | tcagggtcgc  | aggggcaacg  | tgtattgggt | gcgcgactcg | gttagcgccc  | 17580 |
| tgccgctgcc  | cgtgcgcacc  | cgcgcccccg  | gcaactagat | tgcaagaaaa | aactacttag  | 17640 |



|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| actcgtactg  | ttgtatgtat  | ccagcggcgg  | cgccgcgcaa  | cgaagctatg  | tccaagcgca  | 17700 |
| aaatcaaaga  | agagatgctc  | caggtcatcg  | cgccggagat  | ctatggcccc  | ccgaagaagg  | 17760 |
| aagagcagga  | ttacaagccc  | cgaaagctaa  | agcgggtcaa  | aaagaaaaag  | aaagatgatg  | 17820 |
| atgatgaact  | tgacgacgag  | gtggaactgc  | tgacgctac   | cgcccccagg  | cgacgggtac  | 17880 |
| agtggaaaagg | tcgacgcgta  | aaacgtgttt  | tgccacccgg  | caccaccgta  | gtctttacgc  | 17940 |
| ccggtgagcg  | ctccacccgc  | acctacaagc  | gcgtgtatga  | tgaggtgtac  | ggcgacgagg  | 18000 |
| acctgcttga  | gcaggccaac  | gagcgectcg  | gggagtttgc  | ctacggaaaag | cggcataaagg | 18060 |
| acatgctggc  | gttgccgctg  | gacgagggca  | acccaacacc  | tagcctaaag  | cccgtaacac  | 18120 |
| tgccagcagg  | gctgcccgcg  | cttgccaccgt | ccgaagaaaa  | gcgcggccta  | aagcgcgagt  | 18180 |
| ctggtgactt  | ggcaccacc   | gtgcagctga  | tggtacccaa  | gcgccagcga  | ctggaagatg  | 18240 |
| tcttgaaaaa  | aatgaccgtg  | gaacctgggc  | tgagccccga  | gggtccgctg  | cgccaatca   | 18300 |
| agcaggtggc  | gcggggactg  | ggcgtgcaga  | ccgtggacgt  | tcagataccc  | actaccagta  | 18360 |
| gcaccagtat  | tgccaccgcc  | acagagggca  | tgagagacaca | aacgtccccg  | gttgccctcag | 18420 |
| cggtggcgga  | tgccgcgggtg | caggcggtcg  | ctgcccgcgc  | gtccaagacc  | tctacggagg  | 18480 |
| tgcaaacgga  | cccgtggatg  | tttcgcgttt  | cagccccccg  | gcgcccgcg   | ggttcgagg   | 18540 |
| agtacggcgc  | cgccagcgcg  | ctactgcccg  | aatatgccct  | acatccttcc  | attgccccta  | 18600 |
| cccccggtta  | tcgtggctac  | acctaccgcc  | ccagaagacg  | agcaactacc  | cgacgccgaa  | 18660 |
| ccaccactgg  | aacccgcgcg  | cgccgtcgcc  | gtcgccagcc  | cgtgctggcc  | ccgatttccg  | 18720 |
| tgccgagggt  | ggctcgcgaa  | ggaggcagga  | ccctgggtgct | gccaacagcg  | cgctaccacc  | 18780 |
| ccagcatcgt  | gttaaaagccg | gtctttgtgg  | ttcttgacaga | tatggccctc  | acctgccgcc  | 18840 |
| tccgtttccc  | ggtgccggga  | ttccgaggaa  | gaatgacccg  | taggaggggg  | atggccggcc  | 18900 |
| acggcctgac  | gggcggcatg  | cgtcgtgcgc  | accaccggcg  | gcggcgcgcg  | tcgcaccgtc  | 18960 |
| gcattgcgcg  | cggtatcctg  | cccctcctta  | ttccactgat  | cgccgcggcg  | attggcgccg  | 19020 |
| tgcccggaat  | tgcatccgtg  | gccttgacag  | cgacagagaca | ctgattaaaa  | acaagttgca  | 19080 |
| tgtggaaaaa  | tcaaaataaaa | aagtctggac  | tctcacgctc  | gcttggtcct  | gtaactattt  | 19140 |
| tgtagaatgg  | aagacatcaa  | ctttgcgtct  | ctggccccgc  | gacacggctc  | gcgcccgctt  | 19200 |
| atgggaaaact | ggcaagatat  | cggcaccagc  | aatatgagcg  | gtggcgccct  | cagctggggc  | 19260 |
| tcgctgtgga  | gcggcattaa  | aaatctcggt  | tccaccgtta  | agaactatgg  | cagcaaggcc  | 19320 |
| tggaacagca  | gcacaggcca  | gatgctgagg  | gataagttga  | aagagcaaaa  | tttccaacaa  | 19380 |
| aaggtggtag  | atggcctggc  | ctctggcatt  | agcgggggtg  | tggaacctggc | caaccaggca  | 19440 |
| gtgcaaaaata | agattaacag  | taagcttgat  | ccccgccctc  | ccgtagaggga | gcctccaccg  | 19500 |
| gccgtggaga  | cagtgtctcc  | agagggggcg  | ggcgaaaaagc | gtccgcgccc  | cgacagggaa  | 19560 |
| gaaactctgg  | tgacgcaaat  | agacgagcct  | ccctcgtagc  | aggaggcact  | aaagcaaggc  | 19620 |
| ctgcccacca  | cccgtcccat  | cgcccccattg | gctaccggag  | tgctggggcca | gcacacaccc  | 19680 |
| gtaacgctgg  | acctgcctcc  | ccccgccgac  | accagcagga  | aacctgtgct  | ggcaggcccg  | 19740 |
| accgcccgtt  | ttgtaacccg  | tcctagccgc  | gcgtccctgc  | gcccgcgcgc  | cagcgggtccg | 19800 |
| cgatcgttgc  | ggcccgtagc  | cagtggcaac  | tggaacagca  | caactgaacag | catcggtgggt | 19860 |
| ctgggggtgc  | aatccctgaa  | gcgcgcagca  | tgcttctgaa  | tagctaactgt | gtcgtatgtg  | 19920 |
| tgcatgtat   | gcgtccatgt  | cgccgccaga  | ggagctgctg  | agccgcccg   | cgcccgcttt  | 19980 |
| ccaagatggc  | taccctctcg  | atgatgccgc  | agtgggttta  | catgcacatc  | tcggggccagg | 20040 |
| acgcctcgga  | gtacctgagc  | cccgggttgg  | tgacgtttgc  | ccgcgccacc  | gagacgtact  | 20100 |
| tcagcctgaa  | taacaagttt  | agaaacccca  | cggtggcgcc  | tacgcacgac  | gtgaccacag  | 20160 |
| accggtccca  | gcgtttgacg  | ctgcggttca  | tccttggtga  | ccgtgaggat  | actgcgtact  | 20220 |
| cgtacaaggc  | gcggttcacc  | ctagctgtgg  | gtgataaccg  | tggtctggac  | atggcttcca  | 20280 |
| cgtactttga  | catccgcggc  | gtgctggaca  | ggggccctac  | ttttaagccc  | tactctggca  | 20340 |
| ctgcctacaa  | cgccctggct  | cccaagggtg  | ccccaaatcc  | ttgcgaatgg  | gatgaagctg  | 20400 |
| ctactgctct  | tgaaataaac  | ctagaagaag  | aggacgatga  | caacgaagac  | gaagtagacg  | 20460 |
| agcaagctga  | gcagcaaaaa  | actcacgtat  | ttgggcaggc  | gccttattct  | ggtataaata  | 20520 |
| ttacaaagga  | gggtattcaa  | ataggtgtcg  | aaggtcaaac  | acctaaatat  | gccgataaaa  | 20580 |
| catttcaacc  | tgaacctcaa  | ataggagaat  | ctcagtggtg  | cgaaactgaa  | attaatcatg  | 20640 |
| cagctggggag | agtccttaaa  | aagactacc   | caatgaaacc  | atgttacggt  | tcatatgcaa  | 20700 |
| aaaccacaaa  | tgaaaatgga  | gggcaaggca  | ttcttgtaaa  | gcaacaaaat  | ggaaagctag  | 20760 |
| aaagtcaagt  | ggaaatgcaa  | tttttctcaa  | ctactgaggc  | gaccgcaggc  | aatgggtgata | 20820 |
| acttgactcc  | taaagtggta  | ttgtacagtg  | aagatgtaga  | tatagaaacc  | ccagacactc  | 20880 |
| atatttctta  | catgcccact  | attaagggaag | gtaactcacg  | agaactaatg  | ggccaacaat  | 20940 |
| ctatgcccac  | caggccta    | tacattgctt  | ttagggacaa  | ttttattgggt | ctaattgtatt | 21000 |
| acaacagcac  | gggtaatatg  | ggtgttctgg  | cgggccaagc  | atcgagttg   | aatgctgttg  | 21060 |
| tagatttgca  | agacagaaac  | acagagcttt  | cataccagct  | tttgcttgat  | tccattgggtg | 21120 |
| atagaaccag  | gtacttttct  | atgtggaatc  | aggctgttga  | cagctatgat  | ccagatgtta  | 21180 |
| gaattattga  | aaatcatgga  | actgaagatg  | aacttccaaa  | ttactgcttt  | ccactggggag | 21240 |
| gtgtgattaa  | tacagagact  | cttaccgaag  | taaaacctaa  | aacaggtcag  | gaaaatggat  | 21300 |
| gggaaaaaga  | tgctacagaa  | ttttcagata  | aaaatgaaat  | aagagttgga  | aataattttg  | 21360 |



|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccatggaaat  | caatctaaat  | gccaaacctgt | ggagaaat    | cctgtactcc  | aacatagcgc  | 21420 |
| tgtatttgcc  | cgacaagcta  | aagtacagtc  | cttccaacgt  | aaaaatttct  | gataacccaa  | 21480 |
| acacctacga  | ctacatgaac  | aagcgagtg   | tggctcccgg  | gttagtggac  | tgctacatta  | 21540 |
| accttgagc   | acgctgggtcc | cttgactata  | tggacaacgt  | caaccattt   | aaccaccacc  | 21600 |
| gcaatgctgg  | cctgcgctac  | cgctcaatgt  | tgctgggcaa  | tggtcgctat  | gtgcccttcc  | 21660 |
| acatccaggt  | gcctcagaag  | ttctttgcca  | ttaaaaacct  | ccttctcctg  | ccgggctcat  | 21720 |
| acacctacga  | gtggaacttc  | aggaaggatg  | ttaacatggt  | tctgcagagc  | tccttaggaa  | 21780 |
| atgacctaa   | ggttgacgga  | gccagcatta  | agtttgatag  | catttgctt   | tacgccacct  | 21840 |
| tcttccccat  | ggcccaaac   | accgcctcca  | cgcttgaggc  | catgcttaga  | aacgacacca  | 21900 |
| acgaccagtc  | ctttaacgac  | tatctctccg  | ccgccaacat  | gctctaccct  | ataccgcca   | 21960 |
| acgctaccaa  | cgtgcccata  | tccatcccct  | cccgcaactg  | ggcggcttcc  | cgcggtggg   | 22020 |
| ccttcacgcg  | ccttaagact  | aaggaaaccc  | catcactggg  | ctcgggctac  | gacccttatt  | 22080 |
| acacctactc  | tggctctata  | ccctacctag  | atggaacctt  | ttacctcaac  | cacaccttta  | 22140 |
| agaagggtggc | cattaccttt  | gactcttctg  | tcagctggcc  | tggcaatgac  | cgcttgctta  | 22200 |
| cccccaacga  | gtttgaaatt  | aagcgctcag  | ttgacgggga  | gggttacaac  | gttgccaggt  | 22260 |
| gtaacatgac  | caaagactgg  | ttcctgggtac | aaatgctagc  | taactacaac  | attggctacc  | 22320 |
| agggcttcta  | tatcccagag  | agctacaagg  | accgcatgta  | ctccttcttt  | agaaacttcc  | 22380 |
| agcccatgag  | ccgtcagggtg | gtggatgata  | ctaaatacaa  | ggactaccaa  | caggtgggca  | 22440 |
| tcctacacca  | acacaacaac  | tctggatttg  | ttggctacct  | tgccccacc   | atgcgcgaag  | 22500 |
| gacaggccta  | ccctgctaac  | ttcccctatc  | cgcttatagg  | caagaccgca  | gttgacagca  | 22560 |
| ttaccacagaa | aaagtttctt  | tgcgatcgca  | ccctttggcg  | catcccatcc  | tccagtaact  | 22620 |
| ttatgtccat  | gggcgactc   | acagacctgg  | gccaaaacct  | tctctacgcc  | aactccgccc  | 22680 |
| acgcgctaga  | catgactttt  | gaggtggatc  | ccatggacga  | gcccaacctt  | ctttatgttt  | 22740 |
| tggttgaa    | ctttgacgtg  | gtccgtgtgc  | accggccgca  | ccgcggcgctc | atcgaaaccg  | 22800 |
| tgtacctgcg  | cacgcccctt  | tcggccggga  | acgcacacaac | ataaagaagc  | aagcaacatc  | 22860 |
| aacaacagtc  | gccgccatgg  | gctccagtga  | gcaggaactg  | aaagccattg  | tcaaagatct  | 22920 |
| tggttggtggg | ccatattttt  | tgggcacctc  | tgacaagcgc  | tttccagggt  | ttgtttctcc  | 22980 |
| acacaagctc  | gcctgcgcca  | tagtcaatac  | ggccgggtcgc | gagactgggg  | gcgtacactg  | 23040 |
| gatggccttt  | gcctggaacc  | cgactcaaaa  | aacatgctac  | ctctttgagc  | cccttggtct  | 23100 |
| ttctgaccag  | cgactcaagc  | aggtttacca  | gtttgagtac  | gagtcactcc  | tgccgcgtag  | 23160 |
| cgccattgct  | tcttcccccg  | accgctgtat  | aacgctggaa  | aagtccaccc  | aaagcgtaca  | 23220 |
| ggggcccaac  | tcggccgcct  | gtggactatt  | ctgctgcagt  | tttctccacg  | cctttgccaa  | 23280 |
| ctggccccaa  | actcccatgg  | atcacaaacc  | cacctgaac   | cttattaccg  | gggtacccaa  | 23340 |
| ctccatgctc  | aacagtcccc  | aggtacagcc  | caccctgcgt  | cgcaaccagg  | aacagctcta  | 23400 |
| cagcttccctg | gagcgccact  | cgccctactt  | ccgcagccac  | agtgcgcaga  | ttaggagcgc  | 23460 |
| cacttctttt  | tgctacttga  | aaaacatgta  | aaaataatgt  | actagagaca  | ctttcaataa  | 23520 |
| aggcaaatgc  | ttttatttgg  | acactctcgg  | gtgattattt  | acccccaccc  | ttgccgtctg  | 23580 |
| cgccgtttta  | aaatcaaagg  | ggttctgccc  | cgcatcgcta  | tgccgcaactg | gcagggacac  | 23640 |
| gttgcgatac  | tggtggttag  | tgctccactt  | aaactcaggg  | acaaccatcc  | ggcgagctc   | 23700 |
| ggtgaagttt  | tcactccaca  | ggctgcgcac  | catcaccaac  | gcgtttagca  | gctcgggcgc  | 23760 |
| cgatatcttg  | aagtgcgagt  | tggggcctcc  | gccctgcgcg  | cgcgagttgc  | gatacacagg  | 23820 |
| gttgcgacac  | tggaacacta  | tcagcgccgg  | gtgggtgcacg | ctggccagca  | cgctcttgct  | 23880 |
| ggagatcaga  | tcgcggtcca  | ggtccctcgc  | gttgcctcagg | gcgaacggag  | tcaactttgg  | 23940 |
| tagctgcctt  | ccccaaaagg  | gcgcgtgccc  | aggctttgag  | ttgcactcgc  | accgtagtgg  | 24000 |
| catcaaaagg  | tgaccgtgcc  | cggtctgggc  | gttaggatac  | agcgccctgca | taaaagcctt  | 24060 |
| gatctgctta  | aaagccacct  | gagcctttgc  | gccttcagag  | aagaacatgc  | cgcaagactt  | 24120 |
| gccggaaaaac | tgattggccg  | gacaggccgc  | gtcgtgcacg  | cagcaccttg  | cgctcggtgt  | 24180 |
| ggagatctgc  | accacatttc  | ggccccaccg  | gttcttcacg  | atcttggcct  | tgctagactg  | 24240 |
| ctccttcagc  | gcgcgctgcc  | cgttttcget  | cgtcacatcc  | atttcaatca  | cgtgctcctt  | 24300 |
| atztatcata  | atgcttccgt  | gtagacactt  | aagctcgctt  | tcgatctcag  | cgcagcggtg  | 24360 |
| cagccacaac  | gcgcagcccg  | tgggctcgtg  | atgcttgtag  | gtcacctctg  | caaacgactg  | 24420 |
| caggtagccc  | tgagggaatc  | gccccatcat  | cgctacaaag  | gtcttggtgc  | tggtgaagggt | 24480 |
| cagctgcaac  | ccgcggtgct  | cctcgttcag  | ccaggctctg  | catacgcccg  | ccagagcttc  | 24540 |
| cacttggtca  | ggcagtagtt  | tgaagttcgc  | ctttagatcg  | ttatccacgt  | ggtagcttgc  | 24600 |
| catcagcgcg  | cgcgagcctt  | ccatgccttc  | ctcccacgca  | gacacgatcg  | gcacactcag  | 24660 |
| cggttctcat  | accgtaattt  | cactttccgc  | ttcgttgggc  | tcttctctct  | cctcttgctg  | 24720 |
| ccgcatacca  | cgcgccactg  | ggctcgtctt  | attcagccgc  | cgcactgtgc  | gcttacctcc  | 24780 |
| tttgccattgc | ttgattagca  | ccggtgggtt  | gctgaaaccc  | accatttcta  | gcgccacatc  | 24840 |
| ttctctttct  | tcctcgctgt  | ccacgattac  | ctctggtgat  | ggcgggcgct  | cggtcttggg  | 24900 |
| agaaggggcg  | ttctttttct  | tcttggggcg  | aatggccaaa  | tccgcccgcg  | aggtcgatgg  | 24960 |
| ccgcgggctg  | ggtgtgcgcg  | gcaccagcgc  | gtcttgtgat  | gagttctcct  | cgctcctgga  | 25020 |
| ctcgatacgc  | cgcctcatcc  | gcttttttgg  | ggcgcccggg  | ggaggcgggc  | gcgacgggga  | 25080 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| cggggacgac  | acgtcctcca  | tggttggggg  | acgtcgcgcc  | gcaccgcgtc  | cgcgctcggg  | 25140 |
| ggtggtttcg  | cgctgctcct  | cttcccgcact | ggccattttcc | ttctcctata  | ggcagaaaaa  | 25200 |
| gatcatggag  | tcagtcgaga  | agaaggacag  | cctaaccgcc  | ccctctgagt  | tcgccaccac  | 25260 |
| cgcctccacc  | gatgccgcca  | acgcgcctac  | caccttcccc  | gtcagaggcac | ccccgcttga  | 25320 |
| ggaggaggaa  | gtgattatcg  | agcaggaccc  | agggttttgta | agcgaagacg  | acgaggaccg  | 25380 |
| ctcagtagca  | acagaggata  | aaaagcaaga  | ccaggacaac  | gcagaggcaa  | acgaggaaaca | 25440 |
| agtcggggcg  | ggggacgaaa  | ggcatggcga  | ctacctagat  | gtgggagacg  | acgtgctgtt  | 25500 |
| gaagcatctg  | cagcgccagt  | gcgccattat  | ctgcgacgcg  | ttgcaagagc  | gcagcgatgt  | 25560 |
| gccccctgcc  | atagcggatg  | tcagccttgc  | ctacgaacgc  | cacctattct  | caccgcgcgt  | 25620 |
| acccccaaaa  | cgccaagaaa  | acggcacatg  | cgagcccaac  | ccgcgcctca  | acttctaccc  | 25680 |
| cgtattttgcc | gtgccagagg  | tgcttgccac  | ctatcacatc  | tttttccaaa  | actgcaagat  | 25740 |
| acccctatcc  | tgccgtgcca  | accgcagccg  | agcgggacaag | cagctggcct  | tgccggcaggg | 25800 |
| cgtgtgcata  | cctgatatcg  | cctcgctcaa  | cgaagtgcga  | aaaatcttgg  | agggctctgg  | 25860 |
| acgcgcgag   | aagcgcgcg   | caaacgcctc  | cgaaacaggaa | aacagcgaaa  | atgaaagtca  | 25920 |
| ctctggagtg  | ttggtggaac  | tcgagggtga  | caacgcgcgc  | ctagccgtac  | taaaacgcag  | 25980 |
| catcgaggtc  | acccactttg  | cctaccggcg  | acttaacctt  | ccccccaagg  | tcatgagcac  | 26040 |
| agtcatgagt  | gagctgatcg  | tgccgcgtgc  | gcagcccttg  | gagagggatg  | caaatttgca  | 26100 |
| agaacaaaaca | gaggaggggc  | taccgcgagt  | tgccgacgag  | cagctagcgc  | gctggcttca  | 26160 |
| aacgcgcgag  | cctgccgact  | tgaggagcgc  | acgcaaaact  | atgatggccg  | cagtgcctgt  | 26220 |
| taccgtggag  | cttgagtga   | tgacgcggtt  | ctttgctgac  | ccggagatgc  | agcgaagct   | 26280 |
| agaggaaaca  | ttgcactaca  | cctttcgaca  | gggctacgta  | cgccaggcct  | gcaagatctc  | 26340 |
| caacgtggag  | ctctgcaacc  | tggtctccta  | ccttggaatt  | ttgcacgaaa  | accgccttgg  | 26400 |
| gcaaaacgtg  | cttcattcca  | cgctcaaggg  | cgaggcgcg   | cgcgactacg  | tccgcgactg  | 26460 |
| cgtttactta  | tttctatgct  | acacctggca  | gacggccatg  | ggcgtttggc  | agcagtgctt  | 26520 |
| ggaggagtg   | aacctcaagg  | agctgcagaa  | actgctaaa   | caaaacttga  | aggactatg   | 26580 |
| gacggccttc  | aacgagcgct  | ccgtggccgc  | gcacctggcg  | gacatcattt  | tccccgaacg  | 26640 |
| cctgcttaaa  | acccctgaac  | agggctctgc  | agacttcacc  | agtcaaaagca | tgttgacaga  | 26700 |
| ctttaggaac  | ttatctctag  | agcgtcagg   | aatcttgccc  | gccacctgct  | gtgcacttcc  | 26760 |
| tagcgacttt  | gtgcccatta  | agtaccgcga  | atgcccctcg  | ccgctttggg  | gccactgcta  | 26820 |
| ccttctgcag  | ctagccaact  | accttgccct  | ccactctgac  | ataatggaag  | acgtgagcgg  | 26880 |
| tgacggtcta  | ctggagtgtc  | actgtcgtg   | caacctatgc  | accccgacc   | gctccctggt  | 26940 |
| ttgcaattcg  | cagctgctta  | acgaaagtca  | aattatcggt  | acctttgagc  | tgccagggtc  | 27000 |
| ctgcctgac   | gaaaagtccg  | cggtccggg   | gttgaaactc  | actccggggc  | tgtggacgtc  | 27060 |
| ggcttacctt  | cgcaaatctg  | tacctgagga  | ctaccacgcc  | cacgagatta  | ggttctacga  | 27120 |
| agaccaatcc  | cgcccgcaca  | atgcggagct  | taccgcctgc  | gtcattacc   | agggccacat  | 27180 |
| tcttgggcaa  | ttgcaagcca  | tcaacaaagc  | ccgccaagag  | tttctgctac  | gaaaaggacg  | 27240 |
| gggggtttac  | ttggaccccc  | agtccggcga  | ggagctcaac  | ccaatcccc   | cgccgcgcga  | 27300 |
| gccctatcag  | cagcagccgc  | gggcccctgc  | ttcccaggat  | ggcaccacaa  | aagaagctgc  | 27360 |
| agctgcccgc  | gccacccaag  | gacgaggagg  | aatactggga  | cagtacaggca | gaggaggttt  | 27420 |
| tggaacgagga | ggaggaggac  | atgatggaag  | actgggagag  | cctagacgag  | gaagcttccg  | 27480 |
| aggtcgaaga  | ggtgtcagac  | gaaacaccgt  | cacctcgggt  | cgcttcccc   | tcgcccggcg  | 27540 |
| cccagaaatc  | ggcaaccggg  | tccagcatgg  | ctacaacctc  | cgctcctcag  | gcgcgcgcgg  | 27600 |
| cactgcccgt  | tcgcccgaacc | aaccgtagat  | gggacaccac  | tggaaccagg  | gccggtaagt  | 27660 |
| ccaagcagcc  | gcccgcgtta  | gcccgaagag  | aacaacagcg  | ccaaggctac  | cgctcatggc  | 27720 |
| gcgggcacaa  | gaacgccata  | gttgcttgct  | tgcaagactg  | tgggggcaac  | atctccttcg  | 27780 |
| cccgcgcgtt  | tcttctctac  | catcacggcg  | tgcccttccc  | ccgtaacatc  | ctgcattact  | 27840 |
| accgtcatct  | ctacagccca  | tactgcaccg  | gcggcagcgg  | cagcggcagc  | aacagcagcg  | 27900 |
| gccacacaga  | agcaaaaggcg | accggatagc  | aagactctga  | caaagcccaa  | gaaatccaca  | 27960 |
| gcggcgagca  | cagcaggagg  | aggagcgtg   | cgtctggcgc  | ccaacgaacc  | cgtatcgacc  | 28020 |
| cgcgagctta  | gaaacaggat  | ttttccact   | ctgtatgcta  | tatttcaaca  | gagcaggggc  | 28080 |
| caagaacaag  | agctgaaaa   | aaaaaacagg  | tctctgcat   | ccctcaccgg  | cagctgcctg  | 28140 |
| tatcacaaaa  | gcgaagatca  | gcttcggcgc  | acgctggaag  | acgcggaggc  | tctcttcagt  | 28200 |
| aaatactgcg  | cgctgactct  | taaggactag  | tttcgcgcgc  | ttctcfaat   | ttaagcgcga  | 28260 |
| aaactacgtc  | atctccaggc  | ggcagacccg  | ggggcagcac  | ctatccatcag | ccgcattatg  | 28320 |

ttcacgcctc gtcaggcaat cctaactctg cagacctcgt cctctgagcc gcgctctgga 28860  
ggcattggaa ctctgcaatt tattgaggag tttgtgccat cgggtctactt taacccttc 28920  
tcggggacct cgggccaacta tccggatcaa tttattccta actttgacgc ggtaaaggac 28980  
tcggcggaag gctacgactg aatgttaagt ggagaggcag agcaactgcg cctgaaacac 29040  
ctgggtccact gtcgcccga caagtgtttt gcccgcgact ccggtgagtt ttgctacttt 29100  
gaattgcccc aggatcatat cgagggcccc ggcgacggcg tccggcttac cgcccaggga 29160  
gagcttgccc gtgacctgat tcgggagttt aactgtccta accttggtt acatcaagat 29280  
aggggacct gtgttctcac ctacactttt ttcatacatt gcccaagaat aaagaatcgt 29340  
ttaattaatt gccacatcct tttatttttt aattgcagaa aatttcaagt catttttcat 29400  
ttgtgttatg tttcaacgtg ccacatagct cctcccaaca cacagagtac aatcaaaactc 29460  
tcagtagtat agccccacca ctgccacctc cctcccaaca cacagagtac acagtccttt 29520  
acagaacctt agtattcaac agcatcatat catgggtaac agacatattc ttaggtgtta 29580  
ctccccggct ggccttaaaa cgagccaac gctcatcagt gatattaata aactccccg 29640  
tattccacac ggtttcctgt tcgctgtcca gctgctgagc cacaggctgc tgtccaactt 29700  
gcagctcact taagttcatg gaaggagaag tccacgccta catgggggta gagtataat 29760  
gcggttgctt aacgggcgcc tggtgctgca gcagcgcgcg aataaaactgc tgccgcccgc 29820  
cgtgcatcag gataggcgcg gcagcgcgcg agcgatgatt agcgatgatt cgcacggccc 29880  
gctccgctct gcaggaatac aacatggcag agcagcgcac cctgatctca cttaaatcag 29940  
gcagcataag gcgccttgct accacaatat ccacagtgcc cccacagtgcc aaggcgctgt 30000  
cacagtaact catggcgggg accacagaac ccacgtggcc atcataccac aagcgagggt 30060  
atccaaagct gcgaccttc ataaacacgc tggacataaa cattacctt tttggcatgt 30120  
agattaagtg gcgaccttc taccatataa acctctgatt aaacatggcg ccatccacca 30180  
tgtaattcac cactccccg aaaaacctgccc cgccggctat acactgcagg gaaccgggac 30240  
ccatccataa ccagctggcc gcccaggact cgtaacctg gatcatcatg ctcgtcatga 30300  
tggaacaatg acagtggaga ggcacacacg gcatacactt cctcaggatt acaagctcct 30360  
tatcaatgtt aacctatcc cagggaacaa cccattcctg aatcagcgta aatcccacac 30420  
cccgcgttag aacctatcc taactcacgt tgtgcattgt caaagtgtta attcgggca 30480  
tgcagggaag acctcgacg atggtagcgc ggggtttctgt ctcaaaagga ggtagacgat 30540  
gcagcgatg atcctccagt cgagacaacc gagatcgtgt tggtcgtagt gtcagccaa 30600  
ccctactgta cggagtgcgc atatttctg aagcaaaacc aggtgcgggc gtgacaaa 30660  
atggaacgcc ggacgtagtc tccggtctcg cgctctgtgt agtagttgta gtatatccac 30720  
gatctgcgtc tctctcaaag tatccaggc ccccttgct cccggttcta tgtaaaactcc ttcagcgcc 30780  
tctctcaaag tatccaggc ccccttgct cccggttcta tgtaaaactcc ttcagcgcc 30780  
gctgcctga taacatccac caccgcagaa taagccacac ccagccaacc tacacattcg 30840  
ttctgcgagt cacacacggg aggagcggga agagctggaa gaaccatgtt tttttttt 30900  
ttccaaaaga ttaacaaaaa cctcaaaatg aagatctatt aagtgaacgc gctccccctc 30960  
gggtggcgtg tcaaaactca cagccaaaag acagataatg gcatttgtaa gatgttgac 31020  
aatggcttcc aaaaggcaaa cggccctcac gtccaagtgg acgtaaaaggc taacccttc 31080  
agggtgaatc tctctataa acattccagc accttcaacc atgccccaa atttctcatc 31140  
tcgccacct ctcaatata ctctagcaa atcccgaata ttaagtccgg ccattgtaaa 31200  
aatctgctcc agagcgccct ccaccttcag cctcaagcag cgaatcatga ttgcaaaaat 31260  
tcaggttctt cacagacctg tataagattc aaaagcggaa cattaacaaa aataccgcga 31320  
tcccgtaggt ccttcgcag ggccagctga acataatcgt gcaggtctgc acggaccagc 31380  
gcgccactt ccccgccagg aaccttgaca atgtaagctt cactgattat ggttgcatgg 31440  
ctcggaagta tgctaaccag cgtagccccg atgtaagctt tgttgcatgg ggcgcgat 31500  
aaaatgcaag gtgctgctca aaaaatcagg caaagcctcg cgcaaaaag aaagcacatc 31560  
gtagtcatgc tcatgcagat aaaggcaggt aagctccgga accaccacag aaaaagacac 31620  
catttttctc tcaaacatgt ctgcccgtt ctgcataaac acaggaaaaa caaaccttat 31680  
aacatttaaa cattagaagc ctgtcttaca acaggaaaaa caaaccttat aagcataaga 31740  
cggactacgg ccatgccggc gtgaccgtaa aaaaactgggt caccgtgatt aaaaagcacc 31800  
accgacagct cctcggtcat gtcgggagtc ataagttaag actcggtaaa cacatcaggt 31860  
tgattcatcg gtcagtgcta aaaagcgacc gaaatagccc gggggaatac ataccgcag 31920  
gcgtagagac aacattacag cccccatagg aggtataaca aaattaatag gagagaaaaa 31980  
cacataaaca cctgaaaaac cctcctgcct aggcataaata gcaccctccc gctccagaac 32040  
aacatacagc gcttcacagc ggcagcctaa cagtcagcct taccagtaaa aaagaaaacc 32100  
tattaaaaaa acaccactcg acacggcacc agctcaatca gtcacagtggt aaaaaagggc 32160  
caagtgcaga gcgagtatat ataggactaa aacggttaaa gtccacaaaa 32220  
aacaccaga aaaccgcagc cgaacctacg cccagaaaaa aagcacaac 32280  
ttctcaaat cgtcacttcc gttttccac gttacgtaac aagaaaaacta 32340  
caattcccaa cacatacaag ttactccgccc cgtcaccgcg cccgttccca 32400  
cgccccgcgc caagtcacaa actccacccc cttcattatca tattggcttc aatccaaaat 32460  
aaggtatatt attgatgatg 32480

<210> 24  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 24  
 ctcaacaatt gtggatccgt actcc 25

<210> 25  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 25  
 gtgctcagca gatcttgcca ctgtg 25

<210> 26  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 26  
 ggcgcggttcg gatccactct cttcc 25

<210> 27  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 27  
 ctacatgcta ggcagatctc gttcggag 28

<210> 28  
 <211> 1240  
 <212> DNA  
 <213> adenovirus

<400> 28  
 ggatccactc tcttcgcgat cgctgtctgc gagggccagc tgttgggggtg agtactccct 60  
 ctgaaaagcg ggcgatgactt ctgcgctaag attgtcagtt tccaaaaaacg aggaggattt 120  
 gatattcacc tggcccgcgg tgatgccttt gaggggtggc gcatccatct ggtcagaaaa 180  
 gacaatcttt ttgttgtaaa gcttggtggc aaacgacccg tagagggcgt tggacagcaa 240  
 cttggcgatg gagcgacagg tttgggtttt gtcgcgatcg gcgcgctcct tggccgcgat 300  
 gtttagctgc acgtattcgc gcgcaacgca ccgccattcg ggaaagacgg tggcgcgctc 360  
 gtccggcacc aggtgcacgc gccaacgcgc gttgtgcagg gtgacaaggc caacgctggg 420  
 ggctacctct ccgcgtaggc gctcgttggc ccagcagagg cggccgcctt tgcgcgagca 480  
 gaatggcggt aggggggtcta gctgcgtctc gtccgggggg tctgcgtcca cggtaaagac 540  
 cccgggcagc aggcgcgcgc cgaagtagtc tatcttgcac ccttgcaagt ctagcgcttg 600

```

ctgccatgcg cgggcggaag ggcgcgcgtc gtatgggttg agtgggggac cccatggcat 660
gggggtgggtg agcgcgagg aggtacatgcc gcaaatgtcg taaacgtaga ggggctctct 720
gagtatcca agatatgtag ggtagcatct tccaccgcgg atgctggcgc gcacgtaatc 780
gtatagttcg tgcgagggag cgaggagggtc gggaccgagg ttgctacggg cgggctgctc 840
tgctcggaag actatctgcc tgaagatggc atgtgagttg gatgatattg ttggacgctg 900
gaagacgttg aagctggcgt ctgtgagacc taccgcgtca cgcacgaagg aggcgttaga 960
gtcgcgcagc ttgttgacca gctcggcggt gacctgcacg tctagggcgc agtagtccag 1020
ggtttccttg atgatgtcat acttatcctg tccctttttt ttccacagct cgcggttgag 1080
gacaaactct tcgcggtctt tccagtactc ttggatcgga aaccgcgtcg cctccgaacg 1140
agatccgtac tccgcgcgcg agggacctga gcgagtcgc atcgaccgga tcggaaaacc 1200
tctcgagaaa ggcgtctaac cagtacacgt cgcaagatct 1240

```

&lt;210&gt; 29

&lt;211&gt; 8383

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid pDV60

&lt;400&gt; 29

```

gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatagc cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtctat agcccatata 300
tggagttccg cgttacataa cttaacggtaa atggcccgcg tggctgaccg cccaacgacc 360
ccccccatt gagctcaata atgacgtatg aaactgcccc cttggcagta gggactttcc 420
attgacgtca atgggtggac tatttacggt tcaatgacgg taaatggccc gcctggcatt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcggtg acgggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtacc 900
gagctcggat ccactctctt ccgcacgcgt gtctgcgagg gccagctgtt ggggtgagta 960
ctccctctga aaagcgggca tgactttctgc gctaagattg tcagtttcca aaaacgagga 1020
ggatttgata ttcacctggc ccgcggtgat gcctttgagg gtggccgcgt ccactcggtc 1080
agaaaagaca atctttttgt tgtcaagcgt ggtggcaaac gaccgtaga gggcgttgga 1140
cagcaacttg gcgatggagc gcagggtttg gtttttgcg cgatcggcgc gctccttggc 1200
cgcatgtttt agctgcacgt attcgcgcgc aacgcaccgc cattcgggaa agacgggtgtt 1260
gcgctcgtcg ggcaccaggt gcacgcgcca accgcggttg tgcagggtga caaggtcaac 1320
gctggtggct acctctccgc gtaggcgctc gttggtccag cagaggcggc cgcccttgcg 1380
cgagcagaat ggcggtaggg ggtctagctg cgtctcgtcc ggggggtctg cgtccacggt 1440
aaagaccccc cgcgctgctg catgcgcggg gcgctcgtat gtagctctatc ttgcatcctt gcaagctag 1500
tggtcatggg tgggtgagcg cggaggcgta catgccgcaa atgtcgtaaa cgtagagggg 1620
ctctctgagt attccaagat atgtagggtg gcactctcca ccgcggatgc tggcgcgcac 1680
gtaatcgtag agttcgtgcg agggagcgag gaggtcggga ccgaggttgc tacgggcggg 1740
ctgctctgct cggaagacta tctgcctgaa gatggcatgt gagttggatg atatggttgg 1800
acgctggaag acgttgaagc tggcgtctgt gagacctacc gcgtcacgca cgaaggaggc 1860
gtaggagtgc cgcagcttgt tgaccagctc ggcggtgacc tgcacgtcta gggcgagta 1920
gtccaggggt tccttgatga tgtcatactt atcctgtccc ttttttttcc acagctcgcg 1980
gttgaggaca aactcttcgc ggtctttcca gtactcttgg atcgaaaacc cgtcggcctc 2040
cgaacgagat ccgtactccg ccgcccaggg acctgagcga gtccgcacg accggatcgg 2100
aaaaactctc gagaaaggcg tctaaccagt cacagtcgca agatccaaga tgaagcgcgc 2160
aagaccgtct gaagatacct tcaaccccggt gtatccatat gacacggaaa ccggtcctcc 2220
aactgtgcct tttcttactc ctccctttgt atccccaat ggggttcaag agagtcctcc 2280
tgggtgactc tctttgcgcc tatccgaacc tctagtacc tccaatggca tgcttgcgt 2340
caaatgggc aacggcctct ctctggacga gcccggaac cttacctccc aaaatgtaac 2400
cactgtgagc ccacctctca aaaaaaccaa gtcaaacata aacctggaaa tatctgcacc 2460

```

cctcacagtt acctcagaag ccctaactgt ggctgcccgc gcacctctaa tggctcgccgg 2520  
caacacactc accatgcaat cacaggcccc gctaaccgtg cactgactcca aacttagcat 2580  
tgccacccaa ggacccctca cagtgtcaga aggaagctg gccctgcaaa catcaggccc 2640  
cctcaccacc accgatagca gtacccttac tatcactgcc tcacccctc taactactgc 2700  
cactggtagc ttgggcattg acttgaaaga gcccatttat acacaaaatg gaaaactagg 2760  
actaaagtac ggggctcctt tgcatgtaac agacgacctt aacactttga ccgtagcaac 2820  
tgggtccaggt gtgactatta ataatacttc cttgcaaaact aaagtactg gagccttggg 2880  
ttttgattca caaggcaata tgcaacttaa tgtagcagga ggactaagga ttgattctca 2940  
aaacagacgc cttatacttg atgttagtta tccgtttgat gctcaaaacc aactaaatct 3000  
aagactagga cagggccctc tttttataaa ctacagcccac aacttgata ttaactacaa 3060  
caaaggcctt tacttgttta cagcttcaaa caattccaaa aagcttgagg ttaacctaag 3120  
cactgccaag ggggtgatgt ttgacgctac agccatagcc attaatgcag gagatgggct 3180  
tgaatttggg tcacctaattg caccaaacac aaatccctc aaaacaaaaa ttggccatgg 3240  
cctagaattt gattcaaaaca aggctatggg tcctaaacta ggaactggcc ttagtttga 3300  
cagcacaggt gccattacag taggaaacaa aaataatgat aagctaactt tgtggaccac 3360  
accagctcca tctcctaact gtagactaaa tgcagagaaa gatgctaaac tcactttggg 3420  
cttaacaaaaa tgtggcagtc aaataacttg cactagtttca gttttggctg ttaaaggcag 3480  
tttgggtcca atatctggaa cagttcaaaag tccctcatctt attataagat ttgacgaaaa 3540  
tggagtgtca ctaaacaaat ccttccctgga ccagaatat tggaaacttta gaaatggaga 3600  
tcttactgaa ggcacagcct atacaaacgc tgttggattt atgcctaacc tatcagctta 3660  
tccaaaatct cacggtaaaa ctgccaaaag taacattgtc agtcaagttt acttaaacgg 3720  
agacaaaaat aaacctgtaa cactaacctt tacactaaac ggtacacagg aaacaggaga 3780  
cacaactcca agtgcatact ctatgtcatt ttcattgggac tggctctggcc acaactacat 3840  
taatgaaata tttgcccacat cctcttacac ttcttcatatc attgcccag aataaaaaga 3900  
gcggccgctc gagcatgcat ctagagggcc ctattctata gtgtcaccta aatgctagag 3960  
ctcgctgac agcctcgact gtgccttcta gttgccagcc atctgttgtt tgcccctccc 4020  
ccgtgccttc cttgaccctg gaagggtgcca ctcccactgt cctttcctaa taaaatgagg 4080  
aaattgcatc gcattgtctg agtaggtgtc attctattct ggggggtggg gtggggcagg 4140  
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgag gtgggctcta 4200  
tggcttctga ggcggaaaga accagctggg gctctagggg gtatccccac gcgcccgtga 4260  
gcggcgcat t aagcgcgcg ggtgtgtgtg ttacgcgcag cgtgaccgct acacttgcca 4320  
gcgcccagc gcccgctcct ttcgcttctt ctccttcctt ccttgccacg ttccgcccgt 4380  
ttcccgtca agctctaaat cggggcatcc ctttaggggt ccgatttagt gctttacggc 4440  
acctcgacc tcgcccttg acgttggag atgggttcacg tagtgggcca tcgcccgtat 4500  
agacgggttt aacttgaac cctatctcgg cctattcttt taatagtga ctctgttcc 4560  
ggatttcggc ctatttggtta aaaaatgagc tgatttaaca gggattttgg 4620  
tctgtggaat gtgtgtcagt taggggtgtg aaagtccca ggcgtcccag gcaggcagaa 4740  
gtatgcaaa cagcatctc aattagtcag caaccagggt tggaaagtcc ccaggctcc 4800  
cagcaggcag aagtatgcaa agcatgcat tcaattagtc agcaaccata gtcccgcgcc 4860  
taactccgcc catcccgcc ctaactccgc ccagttccgc ccattctccg ccccatggct 4920  
gactaatttt ttttatttat gcagaggccg agggccgctc tgccctctgag ctattccaga 4980  
agttagtgagg aggccttttt ggaggcctag gcttttgcaa aaagctcccg ggagcttgta 5040  
tatccatttt cggatctgat caagagacag gatgaggatc gtttcgcatg attgaacaag 5100  
atggattgca cgcaggttct cccggccgct ggggtggagag gctattcggc tatgactggg 5160  
cacaacagac aatcggtgc tctgatgcc cgggtggag gctgtcagcg caggggcgcc 5220  
cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag 5280  
cgcggtatc gtggctggcc acgacgggcg ttccttgccg agctgtgtc gacgttgtca 5340  
ctgaagcggg aaggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat 5400  
ctcaccttgc tccgtccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata 5460  
cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac 5520  
gtactcggat ggaagccggg cttgtcgatc aggatgatct ggacgaagag catcaggggc 5580  
tcgcgccacc cgaactgttc gccaggctca aggcgcgcac gcccagcggc gaggatctcg 5640  
tcgtgaccca tggcgatgcc tgcttgccga atatcatggg ggaatggc cgcttttctg 5700  
gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggtta 5760  
ccggtgatat tgtgaaagag cttggcggcg aatgggctga ccgcttctc gtgctttacg 5820  
gtatcccgcg tcccgatctg cagcgcacgc ccttctatcg ccttcttgac gagttcttct 5880  
gagcgggact ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga 5940  
tttcgattcc accgccgcct tctatgaaag cttgggcttc ggaatcgttt tccgggacgc 6000  
cggctggatg atcctccagc ggggggatct catgctggag ttcttcgccc accccaactt 6060  
gtttattgca gttacaaata aagcaatagc atcacaaatt tcacaaataa 6120  
agcatttttt tcaactgcatt ctagtgtgtg tttgtccaaa ctcatcaatg tatcttatca 6180

|             |             |             |             |              |             |      |
|-------------|-------------|-------------|-------------|--------------|-------------|------|
| tgtctgtata  | ccgtcgacct  | ctagctagag  | cttggcgtaa  | tcattggatcat | agctgtttcc  | 6240 |
| tgtgtgaaat  | tggtatccgc  | tcacaattcc  | acacaacata  | cgagccggaa   | gcataaagt   | 6300 |
| taaagcctgg  | gggtgcctaat | gagtgagcta  | actcacatta  | attgcgttgc   | gctcactgcc  | 6360 |
| cgctttccag  | tcgggaaacc  | tgctgtgcca  | gctgcattaa  | tgaatcggcc   | aacgcgcggg  | 6420 |
| gagagcggt   | ttgcgtattg  | ggcgctcttc  | cgcttcctcg  | ctcactgact   | cgctgcgctc  | 6480 |
| ggctgttcgg  | ctgcggcgag  | cggtatcagc  | tcactcaaag  | gcggtaatat   | ggttatccac  | 6540 |
| agaatcaggg  | gataacgcag  | gaaagaacat  | gtgagcaaaa  | ggccagcaaa   | aggccaggaa  | 6600 |
| ccgtaaaaag  | gcgcggttgc  | tggtgttttt  | ccataggctc  | cgccccctg    | acgagcatca  | 6660 |
| caaaaatcga  | cgctcaagtc  | agaggtggcg  | aaaccgcaca  | ggactataaa   | gataccaggc  | 6720 |
| gtttccccct  | ggaagctccc  | tcgtgcgctc  | tcctgttccg  | accctgccgc   | ttaccggata  | 6780 |
| cctgtccgcc  | ttctccctt   | cggtgaagcgt | ggcgctttct  | caatgctcac   | gctgtaggta  | 6840 |
| tctcagttcg  | gtgtaggctc  | ttcgctccaa  | gctgggctgt  | gtgcacgaac   | cccccgttca  | 6900 |
| gccccaccgc  | tgccgcttat  | ccgttaacta  | tcgtcttgag  | tccaacccgg   | taagacacga  | 6960 |
| cttatcgcca  | ctggcagcag  | ccactggtaa  | caggattagc  | agagcgaggt   | atgtaggcgg  | 7020 |
| tgctacagag  | ttcttgaagt  | ggtggcctaa  | ctacggctac  | actagaagga   | cagtatttgg  | 7080 |
| tatctgcgct  | ctgctgaagc  | cagttacctt  | cggaaaaaaga | gttggtagct   | cttgatccgg  | 7140 |
| caaacaaacc  | accgctggta  | gcggtgggtt  | ttttgtttgc  | aagcagcaga   | ttacgcgcag  | 7200 |
| aaaaaaagga  | tctcaagaag  | atcctttgat  | ctttctacg   | gggtctgacg   | ctcagtggaa  | 7260 |
| cgaaaaactca | cgtaagggga  | ttttggtcat  | gagattatca  | aaaaggatct   | tcacctagat  | 7320 |
| ccttttaaat  | taaaaatgaa  | gttttaaatc  | aatctaaagt  | atatatgagt   | aaacttggtc  | 7380 |
| tgacagttac  | caatgcttaa  | tcagtgggac  | acctatctca  | gcgatctgtc   | tatttcgttc  | 7440 |
| atccatagtt  | gcctgactcc  | ccgtcgtgta  | gataactacg  | atacgggagg   | gcttaccatc  | 7500 |
| tggtcccgagt | gctgcaatga  | taccgcgaga  | cccacgctca  | ccggctccag   | atztatcagc  | 7560 |
| aataaaccag  | ccagccggaa  | gggcccagcg  | cagaagtggg  | cctgcaactt   | tatccgcctc  | 7620 |
| catccagttc  | attaattggt  | gccgggaagc  | tagagtaaagt | agttcgccag   | ttaatagttt  | 7680 |
| gcgcaacggt  | gttgccattg  | ctacaggcat  | cgtgggtgta  | cgctcgtcgt   | ttgggtatggc | 7740 |
| ttcattcagc  | tccggttccc  | aacgatcaag  | gcgagttaca  | tgatccccca   | tggtgtgcaa  | 7800 |
| aaaagcgggt  | agctccttcg  | gtcctccgat  | cgttgtcaga  | agtaagttgg   | ccgcagtgtt  | 7860 |
| atcactcatg  | gttatggcag  | cactgcataa  | ttccttact   | gtcatgccat   | ccgtaagatg  | 7920 |
| cttttctgtg  | actggtgagt  | actcaaccaa  | gtcattctga  | gaatagtgtg   | tgccggcgacc | 7980 |
| gagttgctct  | tgcccggcgt  | caatacggga  | taataccgcg  | ccacatagca   | gaactttaaa  | 8040 |
| agtgtctatc  | attggaaaaac | gttcttcggg  | gcgaaaaactc | tcaaggatct   | taccgctgtt  | 8100 |
| gagatccagt  | tcgatgtaac  | ccactcgtgc  | acccaactga  | tcttcagcat   | cttttacttt  | 8160 |
| caccagcggt  | tctgggtgag  | caaaaacagg  | aaggcaaaat  | gcgcgaaaaa   | agggaataag  | 8220 |
| ggcgacacgg  | aaatgttgaa  | tactcatact  | cttccttttt  | caatattatt   | gaagcattta  | 8280 |
| tcagggttat  | tgtctcatga  | gcggatacat  | atttgaatgt  | atttagaaaa   | ataaacaat   | 8340 |
| aggggttccg  | cgcacatttc  | ccgaaaagt   | gccacctgac  | gtc          |             | 8383 |

&lt;210&gt; 30

&lt;211&gt; 7960

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid pDV67

&lt;400&gt; 30

|             |             |             |             |            |            |     |
|-------------|-------------|-------------|-------------|------------|------------|-----|
| gacggatcgg  | gagatctccc  | gatccccctat | ggtcgactct  | cagtacaatc | tgctctgatg | 60  |
| ccgcatagtt  | aagccagtat  | ctgctccctg  | cttgtgtggt  | ggaggtcgct | gagtagtgcg | 120 |
| cgagcaaaat  | ttaagctaca  | acaaggcaag  | gcttgaccga  | caattgcatg | aagaatctgc | 180 |
| ttaggggttag | gcgttttgcg  | ctgcttcgcg  | atgtacgggc  | cagatatacg | cgttgacatt | 240 |
| gattattgac  | tagttattaa  | tagtaatcaa  | ttacgggggtc | attagttcat | agcccatata | 300 |
| tggagttccg  | cgttacataa  | cttacggtaa  | atggcccgcgc | tggtgaccg  | cccaacgacc | 360 |
| cccgccttcc  | gacgtcaata  | atgacgtatg  | ttcccatagt  | aacgccaata | gggactttcc | 420 |
| attgacgtca  | atgggtggag  | tatttacggg  | aaactgccca  | cttggcagta | catcaagtgt | 480 |
| atcatatgcc  | aagtacgccc  | cctattgacg  | tcaatgacgg  | taaatggccc | gcctggcatt | 540 |
| atgcccagta  | catgacctta  | tgggactttc  | ctacttggca  | gtacatctac | gtattagtca | 600 |
| tcgctattac  | catgggtgatg | cggttttggc  | agtacatcaa  | tgggcgtgga | tagcgggttg | 660 |
| actcacgggg  | atttccaagt  | ctccacccca  | ttgacgtcaa  | tgggagtttg | ttttggcacc | 720 |
| aaaatcaacg  | ggactttcca  | aaatgtcgta  | acaactccgc  | cccattgacg | caaatgggcg | 780 |
| gtaggcgtgt  | acgggtgggag | gtctatataa  | gcagagctct  | ctggctaact | agagaacca  | 840 |
| ctgcttactg  | gcttatcgaa  | attaatacga  | ctcactatag  | ggagacccaa | gctggctagc | 900 |



gtttaaactt aagcttggtta ccgagctcgg atccactctc ttccgcatcg ctgtctgcga 960  
gggcccagctg ttgggggtgag tactccctct gaaaagcggg catgacttct gcgctaagat 1020  
tgtcagtttc caaaaacgag gaggatttga tattcacctg gcccgcggtg atgcctttga 1080  
gggtggccgc atccatctgg tcagaaaaga caatcttttt gttgtcaagc ttggtggcaa 1140  
acgacccgta gagggcggtg gacagcaact tggcgatgga gcgcagggtt tggtttttgt 1200  
cgcgatcgcc gcgctccttg gccgcgatgt ttagctgcac gtattcgcg ccaacgcacc 1260  
gccattcggg aaagacgggt gtgcgctcgt cgggcaccag gtgcacgcgc caaccgcggt 1320  
tgtgcagggt gacaagggtca acgctggtgg ctacctctcc gcgtaggcgc tcggttggtcc 1380  
agcagaggcg gccgcccctg cgcgagcaga atggcggtag ggggtctagc tgcgtctcgt 1440  
ccgggggggtc tgcgtccacg gtaaaagacc cgggcagcag gcgcgcgctc aagtagtcta 1500  
tcttgcatcc ttgcaagtct agcgcttctt gccatgcgcg ggcggcaagc gcgcgctcgt 1560  
atgggttgag tgggggaccc catggcatgg ggtgggtgag cgcggaggcg tacatgccgc 1620  
aaatgtcgt aacgtagagg ggctctctga gtattccaag atatgtagg tagcatcttc 1680  
caccgcggat gctgcgcgc gctaatcgt atagttcgtg cgaggagcg aggagtcgg 1740  
gaccgaggtt gctacggggt ggctgctctg ctcggaagac tatctgcctg aagatggcat 1800  
gtgagttgga tgatatggtt ggacgctgga agacgttgaa gctggcgctt gtgagacct 1860  
ccgctcagc cagcaaggag gcgtaggagt cgcgcagctt gttgaccagc gtgagacct 1920  
cctgcacgtc tagggcgagc tagtccaggg tttccttgat gatgtcatac tcggcggtga 1980  
cctttttttt ccacagctcg cgggtgagga caaactcttc gcggtctttc cagtactctt 2040  
ggatcgga aaa cccgtcgcc tccgaacgag atccgtactc cgccgcccag ggacctgagc 2100  
gagtcggcat cgaccggatc gga aaactc tccgaaaagg cgtctaacca gtacagctcg 2160  
caagatccaa gatgaagcgc gcaagaccgt ctgaagatac cttcaacccc gtgtatccat 2220  
atgacacgga aaccggctct ccaactgtgc cttttcttac tctcctctt gtatccccc 2280  
atgggtttca agagagtcct cctgggttac tctctttgcg cctatccgaa cctctagta 2340  
cctccaatgg ctgcttgcg ctcaaaatgg accactgtga gcaacggcct ctctctggac gagggcgca 2400  
accttacctc ccaaaatgta accactgtga gccacacct ctaacaggc ctaacaggc 2460  
taaacctgga aatatctgca cccctcacag ttacctcaga agccctaact gtggctgccg 2520  
ccgcacctct aatggctcgc ggcaacacac taccatgca atcacaggc ccgctaaccg 2580  
tgacgactc caaacttagc attgccacct aaggacctt cacagtgtca gaaggaaagc 2640  
tagccctgca aacatcaggc cccctacca ccaccgatag cagtacctt actatcactg 2700  
cctcaccccc tctaactact gccactggtg gcttgggcat tgacttgaaa gagccattt 2760  
atacacaaaa tggaaaacta ggactaaagt acggggctcc tttgcatgta acagacacc 2820  
taaacacttt gaccgtagca actggtccag gtgtgactat taataatact tccttgcaaa 2880  
ctaaagttag ggttttgatt cacaaggcaa tatgcaactt aatgtagcag 2940  
gaggactaag gattgattct caaaacagac gccttatact tgatgttagt tatccggtt 3000  
atgctcaaaa ctaactaaat ctaagactag gacagggccc tctttttata aactcagccc 3060  
acaacttggga tattaactac acaaaaggcc tttacttggt tacagcttca aacaattcca 3120  
aaaagcttga ggttaacct aagcactgcca aggggttgat gtttgacgct acagccatag 3180  
ccattaatgc agtggccat cttgaatttg gttcacctaa tgcaccaaac acaaatcccc 3240  
tcaaaacaaa aattggccat ggctagaat ttgattcaaa caaggctatg gttcctaaac 3300  
taggaactgg ccttagtttt gacagcacag gtgccattac agtaggaaac aaaaataatg 3360  
ataagctaac tttgtggacc acaccagctc catctcctaa ctgtagacta aatgcagaga 3420  
aagatgctaa actcactttg gtcttaacaa aatgtggcag tcaaatactt gctacagttt 3480  
cagttttggc tgttaaaggc agtttggtc caatatctgg aacagttcaa agtgctcatc 3540  
ttattataag atttgacgaa aatggagtgc tactaaacaa ttcccttctg gaccagaat 3600  
attggaactt tagaaatgga gatcttactg aaggcacagc ctatacaaac gctgttggt 3660  
ttatgcctaa cctatcagct tatccaaaat ctcacggtaa aactgcaaaa agtaacattg 3720  
tcagtcaagt ttacttaaac ggagacaaaa ctaaacctgt aacactaacc attactacta 3780  
acggtacaca ggaacacagga gacacaactc caagtgcata ctctatgtca ttttcatggg 3840  
actggtctgg ccacaactac attaatgaaa tatttgccac atcctcttac actttttcat 3900  
acattgccca agaataaaaag aagcggccgc tctagtggc agccatctgt agggcccggt 3960  
gatcagcctc gactgtgcct tctagtggc agccatctgt tggtttggccc tcccccggtc 4020  
cttccctgac cctggaagggt gccactccca ctgtccttcc ctaataaaaat gaggaaattg 4080  
catcgcatgt tctgagtagg tgtcattcta tctggggggg tgggggtggg caggacagca 4140  
agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt 4200  
ctgagggcga aagaaccagc tggggctcta gggggtatcc ccacgcgccc tgtagcgcg 4260  
cattaagcgc ggccgggtgtg gtggttacgc gcagcgtgac cgtacactt ggcacgccc 4320  
tagcgcccg cctttctgct tcttccctt ccttctcgc caggttcgcc ggctttcccc 4380  
gtcaagctct aaatcggggc atccctttag ggttccgatt tagtgcttta cggcacctcg 4440  
accccaaaaa acttgattag ggtgatgggt cacgtagtgg gccatcgccc tgatagacgg 4500  
tttttcgccc tttgacgtt gagtccactg tctttaatag tggactcttg ttccaaactg 4560  
gaacaacact caacctatc tcggtctatt cttttgattt ataagggatt ttggggattt 4620



cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg 4680  
gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc ccaggcaggc agaagtatgc 4740  
aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtcccaggc tcccagcag 4800  
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 4860  
cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa 4920  
ttttttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc cagaagtagt 4980  
gaggaggctt ttttgagggc ctaggctttt gcaaaaaagct cccgggagct tgtatatcca 5040  
ttttcgatc tgatcagcac gtgttgacaa ttaatcatcg gcatagtata tcggcatagt 5100  
ataatacgac aaggtgagga actaaaccat ggccaagtgt accagtgcg ttcgggtgct 5160  
caccgcgcgc gacgtcgccg gacgcgtcga gttctggacc gaccggctcg ggttctccc 5220  
ggacttcgtg gaggacgact tcgcccgtgt ggtccgggac gacgtgaccc tggtcatcag 5280  
cgcggtccag gaccaggtgg tgcgggacaa caccctggcc tgggtgtggg tgcgcggcct 5340  
ggacgagctg tacgccgagt ggtcggaggt cgtgtccacg aacttccggg acgctccgg 5400  
gccggccatg accgagatcg gcgagcagcc gtggggcgcg gagcaggac tgacacgtgc tgcgcgaccc 5460  
ggccggcaac tcgtggccga atgaaagggt gggcttcgga atcggtttcc gggaacggcg 5520  
cgattccacc gccgccttct ctccagcgcg gggatctcat cactagcatc aaaaatttca ccaacttgtt 5580  
ctggatgac ctccagcgcg gggatctcat cactagcatc aaaaatttca ccaacttgtt 5640  
tattgcagct tataatgggt ggttgggttt gtccaaactc atcaatgtat cttatcatgt 5700  
atttttttca ctgcattcta gctagagctt ggcgtaatca tgggtcatagc tgtttcctgt 5760  
ctgtataccg tcgacctcta caattccaca caacatacga gccggaagca taaagtgtaa 5820  
gtgaaattgt tatccgctca cacttaattt gcgttgcgct cactgcccgc 5880  
agctgggggt ggaacactgt cgtgccagct gctcttccgc ttcctcgctc actgactcgc 5940  
tttccagtgc cgtattgggc cggcgagcgg atacagctca ctcaaaggcg gtaatacggg 6000  
aggcggtttg cggcgagcgg aacgcaggaa cggttttcca taggctccgc cccoctgacg 6060  
atcagggat taaaaaggcc aaatcgacgc tcaagtcaga ggtggcgaat cccgacagga ctataaagat 6120  
taaaaaaggcc aaatcgacgc tcaagtcaga ggtggcgaat cccgacagga ctataaagat 6180  
aaatcgacgc tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6240  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6300  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6360  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6420  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6480  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6540  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6600  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6660  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6720  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6780  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6840  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6900  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 6960  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7020  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7080  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7140  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7200  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7260  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7320  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7380  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7440  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7500  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7560  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7620  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7680  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7740  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7800  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7860  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7920  
tccccctgga agctccctcg tgccgtctcc gaagcgtggc gctttctcaa tgctcacgct 7980

<210> 31  
<211> 30  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: primer

&lt;400&gt; 31

atgggatcca agatgaagcg cgcaagaccg

30

&lt;210&gt; 32

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: primer

&lt;400&gt; 32

cactatagcg gccgcattct cagtcattct

30

&lt;210&gt; 33

&lt;211&gt; 7989

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid pDV69

&lt;400&gt; 33

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gacggatcgg  | gagatctccc  | gatccccctat | ggctcgactct | cagtacaatc  | tgctctgatg  | 60   |
| ccgcatagtt  | aagccagtat  | ctgctccctg  | cttggtgtgt  | ggagggtcgt  | gagtagtgcg  | 120  |
| cgagcaaaat  | ttaagctaca  | acaaggcaag  | gcttgaccga  | caattgcatg  | aagaatctgc  | 180  |
| ttaggggttag | gcgtttttgcg | ctgcttcgcg  | atgtacgggc  | cagatatacg  | cggtgacatt  | 240  |
| gattattgac  | tagttattaa  | tagtaatcaa  | ttacgggggtc | attagttcat  | agcccatata  | 300  |
| tggagttccg  | cgttacataa  | cttacggtaa  | atggccccgcc | tggttgaccg  | cccaacgacc  | 360  |
| cccgcccat   | gacgtcaata  | atgacgtatg  | ttcccatagt  | aacgccaata  | gggactttcc  | 420  |
| attgacgtca  | atgggtggac  | tatttacggg  | aaactgcccc  | cttggcagta  | catcaagtgt  | 480  |
| atcatatgcc  | aagtacggcc  | cctattgacg  | tcaatgacgg  | taaatggccc  | gcctggcatt  | 540  |
| atgccagta   | catgacctta  | tgggactttc  | ctacttggca  | gtacatctac  | gtattagtca  | 600  |
| tcgctattac  | catgggtgatg | cggtttttggc | agtacatcaa  | tggggcgtgga | tagcggtttg  | 660  |
| actcacgggg  | atttccaagt  | ctccacccca  | ttgacgtcaa  | tgggagtttg  | ttttggcacc  | 720  |
| aaaatcaacg  | ggactttcca  | aaatgtcgta  | acaactccgc  | cccattgacg  | caaatgggcg  | 780  |
| gtaggcgtgt  | acgggtgggag | gtctatataa  | gcagagctct  | ctggctaact  | agagaaccca  | 840  |
| ctgcttactg  | gcttatcgaa  | attaatacga  | ctcactatag  | ggagacccaa  | gctggctagc  | 900  |
| gtttaaactt  | aagcttggtg  | ccgagctcgg  | atccactctc  | ttccgcacgc  | ctgtctgcga  | 960  |
| ggggccagctg | ttgggggtgag | tactccctct  | gaaaagcggg  | catgacttct  | gcgctaagat  | 1020 |
| tgctcagttt  | caaaaacgag  | gaggatttga  | tattcacctg  | gcccgcgggtg | atgcctttga  | 1080 |
| gggtggccgc  | atccatcttg  | tcagaaaaga  | caatcttttt  | gttgtcaagc  | ttggtggcaa  | 1140 |
| acgacccgta  | gagggcggtg  | gacagcaact  | tggcgatgga  | gcgcagggtt  | tggtttttgt  | 1200 |
| cgcgatcggc  | gcgctccttg  | gccgcgatgt  | ttagctgcac  | gtattcgcgc  | gcaacgcacc  | 1260 |
| gccattcggg  | aaagacgggtg | gtgcgctcgt  | cgggcaccag  | gtgcacgcgc  | caaccgcggg  | 1320 |
| tgtgcagggt  | gacaagggtca | acgctgggtg  | ctacctctcc  | gcgtaggcgc  | tcgttggtcc  | 1380 |
| agcagaggcg  | gccgccttg   | cgcgagcaga  | atggcggtag  | ggggtctagc  | tgctctcgt   | 1440 |
| ccgggggggtc | tgctccacgc  | gtaaagaccc  | cgggcagcag  | gcgcgcgtcg  | aagtagtcta  | 1500 |
| tcttgcatcc  | ttgcaagtct  | agcgctcgt   | gccatgcgcg  | ggcggcaagc  | gcgcgctcgt  | 1560 |
| atgggttgag  | tgggggaccc  | catggcatgg  | gggtgggtgag | cgcggaggcg  | tacatgccgc  | 1620 |
| aaatgtcgta  | aacgtagagg  | ggctctctga  | gtattccaag  | atatgtaggg  | tagcatcttc  | 1680 |
| caccgcggat  | gctggcgcg   | acgtaatcgt  | atagttcgtg  | cgagggagcg  | aggaggtcgg  | 1740 |
| gaccgaggtt  | gctacggggc  | ggctgctctg  | ctcggaagac  | tatctgcctg  | aagatggcat  | 1800 |
| gtgagttgga  | tgatatggtt  | ggacgctgga  | agacgttgaa  | gctggcgtct  | gtgagacctt  | 1860 |
| ccgcgtcacg  | cacgaaggag  | gcgtaggagt  | cgcgcagctt  | gttgaccagc  | tcggcgggtga | 1920 |
| cctgcacgtc  | tagggcgcg   | tagtccaggg  | tttctctgat  | gatgtcatac  | ttatcctgtc  | 1980 |
| ctttttttt   | ccacagctcg  | cggttgagga  | caaactcttc  | gcggtctttc  | cagtactctt  | 2040 |
| ggatcggaaa  | cccgtcgggc  | tccgaacgag  | atccgtactc  | cgcgcgcgag  | ggacctgagc  | 2100 |
| gagtcgcgat  | cgaccggatc  | ggaaaacctc  | tcgagaaagg  | cgtctaacca  | gtcacagtgc  | 2160 |

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| caagatccaa  | gatgaagcgc  | gcaagaccgt  | ctgaagatac  | cttcaacccc  | gtgtatccat  | 2220 |
| atgacacgga  | aaccggtcct  | ccaactgtgc  | cttttcttac  | tcctcccttt  | gtatccccc   | 2280 |
| atgggtttca  | agagagtccc  | cctgggggtac | tctctttgcg  | cctatccgaa  | cctctagtta  | 2340 |
| cctccaatgg  | catgcttgcg  | ctcaaatgg   | gcaacggcct  | ctctctggac  | gaggccggca  | 2400 |
| accttacctc  | ccaaaatgta  | accactgtga  | gcccacctct  | caaaaaaac   | aagtcaaa    | 2460 |
| taaacctgga  | aatatctgca  | cccctcacag  | ttacctcaga  | agccctaact  | gtggctgcg   | 2520 |
| ccgacacctc  | aatggtcgcg  | ggcaacacac  | tcaccatgca  | atcacaggcc  | ccgctaaccg  | 2580 |
| tgcacgactc  | caaacttagc  | attgccaccc  | aaggacccct  | cacagtgtca  | gaaggaaagc  | 2640 |
| tagccctgca  | aacatcaggc  | cccctcacca  | ccaccgatag  | cagtaccctt  | actatcactg  | 2700 |
| cctcaccccc  | tctaactact  | gccactggta  | gcttgggcat  | tgacttgaaa  | gagccattt   | 2760 |
| atacacaaaa  | tggaaaaacta | ggactaaagt  | acggggctcc  | tttgcatgta  | acagacgacc  | 2820 |
| taaacacttt  | gaccgtagca  | actgggtccag | gtgtgactat  | taataaact   | tccttgcaaa  | 2880 |
| ctaaagttac  | tggagccttg  | ggttttgatt  | cacaaggcaa  | tatgcaactt  | aatgtagcag  | 2940 |
| gaggactaag  | gattgattct  | caaaacagac  | gccttatact  | tgatgttagt  | tatccgtttg  | 3000 |
| atgctcaaaa  | ccaactaaat  | ctaagactag  | gacagggccc  | tctttttata  | aactcagccc  | 3060 |
| acaacttggg  | tattaactac  | aacaaaggcc  | tttacttggt  | tacagcttca  | aacaattcca  | 3120 |
| aaaagcttga  | ggttaacctta | agcactgcca  | aggggttgat  | gtttgacgct  | acagccatag  | 3180 |
| ccattaatgc  | aggagatggg  | cttgaatttg  | gttcaccta   | tgacacaaac  | acaaatcccc  | 3240 |
| tcaaaaacaaa | aattggccat  | ggcctagaat  | ttgattcaaa  | caaggctatg  | gttcctaaac  | 3300 |
| taggaaactgg | ccttagtttt  | gacagcacag  | gtgccattac  | agtaggaaac  | aaaaataatg  | 3360 |
| ataagctaac  | tttgtggacc  | gggtccaaaac | cagaagccaa  | ctgcataatt  | gaatacggga  | 3420 |
| aacaaaaccc  | agatagcaaa  | ctaactttaa  | tccttgtaaa  | aaatggagga  | attgttaatg  | 3480 |
| gatatgtaac  | gctaattggga | gcctcagact  | acgttaacac  | cttattttaa  | aacaaaaatg  | 3540 |
| tctccattaa  | tgtagaacta  | tactttgatg  | ccactggtca  | tatattacca  | gactcatctt  | 3600 |
| ctcttaaaac  | agatctagaa  | ctaaaataca  | agcaaacgcg  | tgactttagt  | gcaagaggtt  | 3660 |
| ttatgccaa   | tactacagcg  | tatccatttg  | tccttcctaa  | tgccgggaaca | cataatgaaa  | 3720 |
| attatatttt  | tgggtcaatgc | tactacaaag  | caagcgatgg  | tgcccttttt  | ccgttggaag  | 3780 |
| ttactgttat  | gcttaataaaa | cgctgccag   | atagtcgcac  | atcctatggt  | atgacttttt  | 3840 |
| tatggctcct  | gaatgctggg  | ctagctccag  | aaactactca  | ggcaaccctc  | ataacctccc  | 3900 |
| catttacctt  | ttcctatatt  | agagaagatg  | actgattttt  | aagaagcggc  | cgctcgagtc  | 3960 |
| tagagggccc  | gtttaaaccc  | gctgatcagc  | ctcgcactgtg | ccttctagtt  | gccagccatc  | 4020 |
| tggtgtttgc  | ccctcccccg  | tgcttccctt  | gaccctggaa  | ggtgccactc  | ccactgtcct  | 4080 |
| ttcctaataa  | aatgaggaaa  | ttgcatcgca  | ttgtctgagt  | aggtgtcatt  | ctattctggg  | 4140 |
| gggtgggggtg | gggcaggaca  | gcaaggggga  | ggattgggaa  | gacaatagca  | ggcatgctgg  | 4200 |
| ggatgcgggtg | ggctctatgg  | cttctgaggc  | ggaaagaacc  | snccntagct  | ggggctctag  | 4260 |
| gggggtatccc | cacgcgccct  | gtagcggcgc  | atlaagcgcg  | gcgggtgtgg  | tggttacgcg  | 4320 |
| cagcgtgacc  | gctacacttg  | ccagcgccct  | agcgcccgct  | cctttcgctt  | tcttcccttc  | 4380 |
| ctttctcgcc  | acgttcgcgc  | gctttccccg  | tcaagctcta  | aatcggggca  | tccttttagg  | 4440 |
| gttccgatttt | agtgttttac  | ggcacctcga  | ccccaaaaaa  | cttgattagg  | gtgatgggtc  | 4500 |
| acgtagtggg  | ccatcgccct  | gatagacggt  | ttttcgccct  | ttgacgttgg  | agtccacggt  | 4560 |
| ctttaaatagt | ggactcttgt  | tccaaactgg  | aacaacactc  | aaccctatct  | cggtctattc  | 4620 |
| ttttgattta  | taagggtattt | tggggatttc  | ggcctattgg  | ttaaaaaatg  | agctgattta  | 4680 |
| acaaaaattt  | aacgcgaatt  | aattctgtgg  | aatgtgtgtc  | agttagggtg  | tggaaagtcc  | 4740 |
| ccaggctccc  | caggcaggca  | gaagtatgca  | aagcatgcat  | ctcaattagt  | cagcaaccag  | 4800 |
| gtgtggaaag  | tccccaggct  | ccccagcagg  | cagaagtatg  | caaagcatgc  | atctcaatta  | 4860 |
| gtcagcaacc  | atagtcccgc  | ccctaactcc  | gcccataccc  | cccctaactc  | cgcccagttc  | 4920 |
| cgcccatctc  | ccgccccatg  | gctgactaat  | tttttttatt  | tatgcagagg  | ccgaggccgc  | 4980 |
| ctctgcctct  | gagctattcc  | agaagttagtg | aggaggcttt  | tttgagggcc  | taggcttttg  | 5040 |
| caaaaagctc  | ccgggagctt  | gtatatccat  | tttcggatct  | gatcagcacg  | tgttgacaat  | 5100 |
| taatcatcgg  | catagtatat  | cggcatagta  | taatacgaca  | aggtgaggaa  | ctaaaccatg  | 5160 |
| gccaagtgtga | ccagtgcctg  | tccggtgctc  | accgcgcgcg  | acgtcgccgg  | agcggctcgag | 5220 |
| ttctggaccg  | accggctcgg  | gttctcccg   | gacttcgtgg  | aggacgactt  | cgccgggtgtg | 5280 |
| gtccgggacg  | acgtgaccct  | gttcatcagc  | gcgggtccagg | accagggtgtg | gccggacaac  | 5340 |
| accctggcct  | gggtgtgggt  | gcgcggcctg  | gacgagctgt  | acgcgagtg   | gtcggagggtc | 5400 |
| gtgtccacga  | acttccggga  | cgctccggg   | ccggccatga  | ccgagatcgg  | cgagcagccg  | 5460 |
| tggggggcggg | agttcgccct  | gcgcgacccg  | gccggcaact  | gcgtgcactt  | cgtggccgag  | 5520 |
| gagcaggact  | gacacgtgct  | acgagatttc  | gattccaccc  | ccgccttcta  | tgaaagggtg  | 5580 |
| ggcttcggaa  | tcgttttccg  | ggacgcgcgg  | tggatgatcc  | tccagcgccg  | ggatctcatg  | 5640 |
| ctggagttct  | tcgcccaccc  | caacttgttt  | attgcagctt  | ataatggtta  | caaataaagc  | 5700 |
| aatagcatca  | caaatttcac  | aaataaagca  | ttttttttcac | tgacttctag  | ttgtgggttg  | 5760 |
| tccaaactca  | tcaatgtatc  | ttatcatgtc  | tgtataccgt  | cgacctctag  | ctagagcttg  | 5820 |
| gcgtaatcat  | ggtcatagct  | gtttccctgtg | tgaaattgtt  | atccgctcac  | aattccacac  | 5880 |

```

aacatacgag ccggaagcat aaagtgtaaa gcctgggggtg cctaagtagt gagctaactc 5940
acattaattg cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg 6000
cattaatgaa tcggccaacg cgcgggggaga ggcggtttgc gtattgggag ctcttccgct 6060
tctcgcgtca ctgactcgct gcgctcggtc gttcgggtgc ggcgagcggt atcagctcac 6120
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 6180
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 6240
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtacagag gtggcgaaac 6300
ccgacaggac tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct 6360
gttccgaccc tgcgcgttac cggatacctg tccgccttcc tcccttcggg aagcggtggc 6420
ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 6480
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 6540
cttgagtcga acccggttaag acacgactta tcgccactgg cagcagccac tggtaacagg 6600
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtgggt gcctaactac 6660
ggctacacta gaaggacagt atttgggtatc tgcgctctgc tgaagccagt taccctcgga 6720
aaaagagttg gtactctctg atccggcaaa caaaccaccg ctggtagcgg tggttttttt 6780
gtttgcaagc agcagattac gtggaacgaa aaaggatctc aagaagatcc tttgatcttt 6840
tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 6900
ttatcaaaaa ggatcttcac ttagtctgac agttaccaat aatgaagttt taaatcaatc 6960
taaagtatat atctcagcga tctgtctatt tcgttcatcc atagtgcct gactccccgt cgtgtagata 7080
actacgatac gggaggggctt accatctggc cccagtgtcg caatgatacc gcgagaccca 7140
cgctcacccg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgaga 7200
agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga 7260
gtaagttagt cgccagttaa tagtttgccg aacgttggtg ccattgctac aggcacgtg 7320
gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga 7380
gttacatgat ccccatggtt gtgcaaaaaa gcggttagct ccttcgggtc tccgatcgtt 7440
gtcagaagta agttggccgc agtggtatca ctcatggtta tggcagcact gcataattct 7500
cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca 7560
ttctgagaat agtgatgcg gcgaccgagt tgccttgcc cggcgtaaat acgggataat 7620
accgcgccac atagcagaac tttaaaagtg gctcatcttg gaaaacgttc ttcggggcga 7680
aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaaccac tctgacccc 7740
aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg 7800
caaaatgccg caaaaaaggg aataaggcg ggttattgtc tcatgagcgg catactcttc 7860
ctttttcaat attattgaag catttatcag gttccgcga catttccccg atacatattt 7920
gaatgtattt agaaaaataa acaaataggg gttccgcga ctttccccg aaaagtgcc 7980
cctgacgtc

```

&lt;210&gt; 34

&lt;211&gt; 7607

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: plasmid GRE5-E1-SV40-Hygro

&lt;400&gt; 34

```

tctagaagat ccgctgtaca ggatgttcta gctactttat tagatccgct gtacaggatg 60
ttctagctac tttattagat ccgctgtaca ggatgttcta gctactttat tagatccgct 120
gtacaggatg ttctagctac tttattagat ccgctgtaca gatgttctag ctactttatt 180
agatcgatct cctggccgtt cgggggtcaaa aaccagggtt ggctataaaa ggggggtggg 240
gcgcttccgt cctcactctc ttccgcacgc ctgtctgcga gggccaggat cgatccctgag 300
aacttcaggg taggtttggg gacccttgat tgttctttct ttttcgctat tgtaaaattc 360
atgtttatat gagggggcaa agttttcagg gtgttggtta gaatgggaag atgtcccttg 420
tatcaccatg gaccctcatg ataattttgt ttctttcact ttctactctg ttgacaacca 480
ttgtctcctc ttattttctt ttcattttct gtaacttttt cgttaaactt tagcttgcat 540
ttgtaacgaa tttttaaatt cactttttgt tatttgteag attgtaagta ctttctctaa 600
tcactttttt ttcaaggcaa tcaggggtata ttatattgta cttcagcaca gttttagaga 660
acaattgtta taattaaatg ataaggtaga atatttctgc atataaattc tggctggcgt 720
ggaaatattc ttattggtag aaacaactac atcctgggtc tcatcctgcc tttctcttta 780
tggttacaat gatatacact gtttgagatg aggataaaat actctgagtc caaaccgggc 840
ccctctgcta accatgttca tgccttcttc tttttctac agctcctggg caactgtctg 900
gttattgtgc tgtctcatca ttttggcaaa gaattagatc taagcttctg cagctcgagg 960

```

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| actcggtcga  | ctgaaaatga  | gacatattat  | ctgccacgga  | gggtgttatta | ccgaagaaat  | 1020 |
| ggccgccagt  | cttttggacc  | agctgatcga  | agaggtagctg | gctgataatc  | ttccacctcc  | 1080 |
| tagccatttt  | gaaccaccta  | cccttcacga  | actgtatgat  | ttagacgtga  | cggcccccga  | 1140 |
| agatcccaac  | gaggaggcgg  | tttcgcagat  | ttttcccgac  | tctgtaatgt  | tggcgggtgca | 1200 |
| ggaagggatt  | gacttactca  | cttttccgcc  | ggcgcccggt  | tctccggagc  | cgcctcacct  | 1260 |
| ttcccggcag  | cccagagcagc | cggagcagag  | agccttgggt  | cgggtttcta  | tgccaaacct  | 1320 |
| tgtaccggag  | gtgatcgatc  | ttacctgcca  | cgaggctggc  | tttccaccca  | gtgacgacga  | 1380 |
| ggatgaagag  | ggtgaggagt  | ttgtgttaga  | ttatgtggag  | caccccgggc  | acgggttgca  | 1440 |
| gtcttgtcat  | tatcaccgga  | ggaatacggg  | ggaccagat   | attatgtgtt  | cgctttgcta  | 1500 |
| tatgaggacc  | tgtggcatgt  | ttgtctacag  | taagtgaata  | ttatgggcag  | tgggtgatag  | 1560 |
| agtgggtgggt | ttgggtgggt  | aatTTTTTTT  | ttaatTTTTT  | cagttttgtg  | gtttaaagaa  | 1620 |
| ttttgtattg  | tgattTTTTT  | aaaaggctct  | gtgtctgaac  | ctgagcctga  | gcccagagcca | 1680 |
| gaaccggagc  | ctgcaagacc  | taccgcgcgt  | cctaaaatgg  | cgcctgctat  | cctgagacgc  | 1740 |
| ccgacatcac  | ctagtgtctag | agaatgcaat  | agtagtacgg  | atagctgtga  | ctccgggtcct | 1800 |
| tctaaccacac | ctcctgagat  | acacccggtg  | gtcccgctgt  | gccccattaa  | accagttgcc  | 1860 |
| gtgagagtgt  | gtgggcgtcg  | ccaggctgtg  | gaatgtatcg  | aggacttgct  | taacgagcct  | 1920 |
| gggcaacctt  | tggacttgag  | ctgtaaacgc  | cccaggccat  | aagggtgtaa  | cctgtgattg  | 1980 |
| cgtgtgtggt  | tacgcctttt  | gtttgctgaa  | tgagttgatg  | taagtTTaat  | aaaggggtgag | 2040 |
| ataatgttta  | acttgcattg  | cgtgttaaat  | ggggcggggc  | ttaaagggta  | tataatgcgc  | 2100 |
| cgtgggctaa  | tcttgggttac | atctgacctc  | atggaggctt  | gggagtgttt  | ggaagatttt  | 2160 |
| tctgtgtgtc  | gtaacttgct  | ggaaacagagc | tctaaccagta | cctcttggtt  | ttggagggtt  | 2220 |
| ctgtggggct  | catcccaggc  | aaagttagtc  | tgcagaatta  | aggaggatta  | caagtgggaa  | 2280 |
| tttgaagagc  | ttttgaaatc  | ctgtgggtgag | ctgtttgatt  | ctttgaaatc  | gggtcaccag  | 2340 |
| gcgcttttcc  | aagagaaggt  | catcaagact  | ttggattttt  | ccacaccggg  | gcgcgtgcgc  | 2400 |
| gctgctgtgt  | cttttttgag  | tattataaag  | gataaatgga  | gcgaagaaac  | ccatctgagc  | 2460 |
| gggggggtacc | tgtctggattt | tctggccatg  | catctgtgga  | gagcgggtgt  | gagacacaag  | 2520 |
| aatcgccctgc | tactgttgtc  | ttccgtccgc  | ccggcgataa  | taccgacgga  | ggagcagcag  | 2580 |
| cagcagcagg  | aggaagccag  | gcggcgggcg  | caggagcaga  | gccccatgga  | cccgagagcc  | 2640 |
| ggcctggacc  | ctcggggaatg | aatgttgtac  | agggtggctga | actgtatcca  | gaactgagac  | 2700 |
| gcatttttgac | aattacagag  | gatgggcagg  | ggctaaaggg  | ggtaaagagg  | gagcgggggg  | 2760 |
| cttgtgaggc  | tacagaggag  | gctaggaatc  | tagcttttag  | cttaatgacc  | agacaccgtc  | 2820 |
| ctgagtgtat  | tacttttcaa  | cagatcaagg  | ataattgcgc  | taatgagctt  | gatctgctgg  | 2880 |
| cgcagaagta  | ttccatagag  | cagctgacca  | cttactggct  | gcagccaggg  | gatgattttg  | 2940 |
| aggaggctat  | tagggatatat | gcaaagggtg  | cacttaggcc  | agattgcaag  | tacaagatca  | 3000 |
| gcaaacttgt  | aaatatcagg  | aatgttgtct  | acatttctgg  | gaacggggcc  | gaggtggaga  | 3060 |
| tagatacggg  | ggatagggtg  | gcctttagat  | gtagcatgat  | aaatatgtgg  | ccgggggtgc  | 3120 |
| ttggcatgga  | cgggggtgggt | attatgaatg  | taagggtttac | tggccccaat  | tttagcggta  | 3180 |
| cggtttttct  | ggccaataacc | aaccttatcc  | tacacgggtg  | aagcttctat  | gggtttaaca  | 3240 |
| atacctgtgt  | ggaagcctgg  | accgatgtaa  | gggttcgggg  | ctgtgccttt  | tactgtgctt  | 3300 |
| ggaaggggtg  | gggtgtgtgc  | cccaaaagca  | gggttcaat   | taagaaatgc  | ctctttgaaa  | 3360 |
| gggtgtacct  | gggtatcctg  | tctgagggtg  | actccagggt  | gcgccacaat  | gtggcctccg  | 3420 |
| actgtgggtg  | cttcattgcta | gtgaaaagcg  | tggctgtgat  | taagcataac  | atggtagtg   | 3480 |
| gcaactgcga  | ggacagggcc  | tctcagatgc  | tgacctgtc   | ggacggcaac  | tgtcacctgc  | 3540 |
| tgaagaccat  | tcacgtagcc  | agccactctc  | gcaaggcctg  | gccagtgttt  | gagcataaca  | 3600 |
| tactgaccgg  | ctgttccttg  | catttgggtg  | acaggagggg  | gggtttccta  | ccttaccat   | 3660 |
| gcaatttgag  | tcacactaag  | atattgcttg  | agcccgagag  | catgtccaag  | gtgaacctga  | 3720 |
| acgggggtgt  | tgacatgacc  | atgaagatct  | ggaagggtgct | gaggtacgat  | gagacccgca  | 3780 |
| ccagggtgcag | accctgcgag  | tgtggcggta  | aacatattag  | gaaccagcct  | gtgatgctgg  | 3840 |
| atgtgaccga  | ggagctgagg  | cccgatcact  | tgggtgctggc | ctgcacccgc  | gctgagtttg  | 3900 |
| gctctagcga  | tgaagataca  | gattgaggta  | ctgaaatgtg  | tgggcgtggc  | ttaggggtgg  | 3960 |
| gaaagaatat  | ataagggtggg | ggtcttatgt  | agttttgtat  | ctgttttgca  | gcagccggccg | 4020 |
| ccgccatgag  | caccaactcg  | tttgatggaa  | gcattgtgag  | ctcatatttg  | acaacgcgca  | 4080 |
| tgcccccatg  | ggccgggggtg | cgtcagaatg  | tgatgggctc  | cagcattgat  | ggtcgccccg  | 4140 |
| tccctgccccg | aaactctact  | accttgacct  | acgagaccgt  | gtctggaacg  | ccgttgagga  | 4200 |
| ctgcagcctc  | cgccgcgcgt  | tcagccgctg  | cagccaccgc  | ccgcgggatt  | gtgactgact  | 4260 |
| ttgcttttct  | gagcccgctt  | gcaagcagtg  | cagcttcccg  | ttcatccgcc  | cgcatgaca   | 4320 |
| agtttacggc  | tcttttggca  | caattggatt  | ctttgaccgg  | ggaacttaat  | gtcgtttctc  | 4380 |
| agcagctggt  | ggatctgcgc  | cagcaggttt  | ctgccctgaa  | ggcttccctc  | cctcccaatg  | 4440 |
| cggtttaaaaa | cataaataaaa | aaaccagact  | ctgtttggat  | ttggatcaag  | caagtgtctt  | 4500 |
| gctgtctcag  | ctgactgctt  | aagtgcgaag  | ccgaattgga  | tccaattcgg  | atcgatctta  | 4560 |
| ttaaagcagg  | acttggttat  | tgacgtttat  | aattggttaca | aataaagcaa  | tagcatcaca  | 4620 |
| aatttcacaa  | ataaagcatt  | tttttactg   | cattctagtt  | gtgggttgtc  | caaactcatc  | 4680 |

```

aatgtatctt atcatgtctg gtcgactcta gactcttccg cttctctcgt cactgactcg 4740
ctgcgctcgg tcgttcgggt gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg 4800
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag 4860
gccaggaacc gtaaaaaggc cgcgttgctg gcggttttcc ataggctccg cccccctgac 4920
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 4980
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 5040
accggatacc tgtccgcctt tctcccttcg ggaagcggtg cggtttctca tagctcacgc 5100
tgtaggtatc tcagttcggt gtaggtcggt cgtcccaagc tgggctgtgt gcacgaaccc 5160
cccgttcagc ccgaccgctg gccttatcc ggtaactatc gtcttgagtc caaccgggta 5220
agacacgact tatcgccact ggcagcagcc actggttaaca ggattagcag agcggaggtat 5280
gtagggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 5340
gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 5400
tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt 5460
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 5520
cagtggaacg aaaactcacg ttaagggaat ttggtcatga gattatcaaa aaggatcttc 5580
acctagatcc acagttacca atgcttaatc agtgaggcac ctatctcagc taactacgat acgggagggc 5640
acttggtaga ccatagttgc ctgactcccc tgcaatgata gccgagcgca gaagtggctc tgcaacttta 5700
tttcggtcat ttttaaatta aaaatgaagt ttaaatcaa tctaaagtat atatgagtaa 5760
ttaccatctg gccccagtg agccggaagg ggtacccctc cactacacac ggctccagat 5820
ttatcagcaa taaccagcc taattgttgc cggggaagcta gagtaagtag ttcgccagtt 5880
tccgctcca tccagtctat tgccattgct cgggttccaa cagatcaaggc tgggtgcacg ctcgctgctt 5940
aatagtttgc gcaacgttgt ctccttcggg cctccgatcg ttgttcacg atccccatg 6000
ggtagggctt cattcagctc tatggcagca ctgcataatt ttgtcagaag taagttggcc 6060
ttgtgcaaaa aagcggttag ctccttcggg tgggtgagca cgtatcaaggc ctcttactgt catgcatcc 6120
gcagtggtat cactcatggt tatggcagca tcaaccaagt ataccgcgcc atagtgatg 6180
gtaagatgct tttctgtgac tgggtgagca cccggcgta ataccgcgcc aaggaatcta 6240
cgcgaccga gttgctcttg tggaaaacgt gatgtaaccc actcgtgcac gaaaactctc ttcagcatct 6300
actttaaaag tgctcatcat gatccagttc tgggtgagca aaaacaggaa ccaactgatc cgcaaaaaag 6360
ccgctgttga ccagcgtttc tgggtgagca atgttgaaata ctcatactct ttgaatgtat ttagaaaaat 6420
tttactttca ccagcgtttc tgggtgagca tctcatgagc ggatacatat cacttgacgt ctaagaaacc 6480
ggaataaagg cgacacggaa aggggttattg ggggttccgc ctataaaaaat agggctatca ctagggcccc 6540
agcatttatc aaacaaatag attattatca tgacattaac aaaacctctg acacatgcag ctcccggaga cgggtcacagc 6600
aaacaaatag ggggttccgc ctataaaaaat agggctatca ctagggcccc ctcccggaga cgggtcacagc 6660
attattatca tgacattaac aaaacctctg acacatgcag ctcccggaga cgggtcacagc 6720
gcgtttcggg gatgacgggt gcggatgccc ggagcagaca atcagagcag attgtactga ggtgcacca 6780
ttgtctgtaa gcggatgccc ggctggctta actatgcggc acagatgcgt aaggagaaaa taccgcatca ggaattgta 6840
cgggtgtcgg gaaataccgc attcgcgtta aatcttctat aatcaaaaag aatagaccga gatagggttg 6900
tatcggtgtg ttttggttaa aatcccttat caagagtcca cctattaaaga acgtggactc caactcaagt 6960
agcgttaata caataggccc cagtttgtaa caagagtcca cctattaaaga acgtggactc caactcaagt 7020
caataggccc cagtttgtaa aatcccttat caagagtcca cctattaaaga acgtggactc caactcaagt 7080
agtgttggtc cagtttgtaa caagagtcca cctattaaaga acgtggactc caactcaagt 7140
gggcgaaaaa ccgtctatca gggcgatggc taaagcacta aatcggaacc ggcgaacgtg ggcgaacgtg 7200
tttttggggt cgaggtgccc ggggaaagcc gggcgctggc aagtgtagcg gtcacgctgc cactcgccat 7260
agagcttgac ggggaaagcc gggcgctggc aagtgtagcg gtcacgctgc cactcgccat 7320
gcggcgctga gggcgctggc aagtgtagcg gtcacgctgc cactcgccat 7380
gcgttaatg cgccgctaca gggcgctggc aagtgtagcg gtcacgctgc cactcgccat 7440
gaagggcgat cggtgcccgc ctcttcgcta ttacgccagc tggcgaaaagg cagcgaggtg taaaacgacg 7500
gcaaggcgat taagttgggt aacgccaggg ttttccagc cagcgaggtg taaaacgacg 7560
gccagtgaaat tgtaatacga ctactatag ggcgaattaa ttcggggg 7607

```

<210> 35  
 <211> 11600  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: plasmid MMTV-E2a-SV40-Neo

<400> 35  
 gaattccgca ttgcagagat attgtattta agtgccatgc tcgatacaat aaacgccatt 60  
 tgaccattca ccacattggt gtgcacctcc aagcttgggc agaaatggtt gaactcccg 120

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gagtgtccta  | cacctagggg  | agaagcagcc  | aaggggttgt  | ttcccaccaa  | ggacgacccg  | 180  |
| tctgcgaca   | aacggatgag  | cccatcagac  | aaagacatat  | tcatttctctg | ctgcaaaactt | 240  |
| ggcatagctc  | tgcttttgct  | ggggctattg  | ggggaagtgt  | cggttcgtgc  | tcgcagggct  | 300  |
| ctcacccttg  | actcttttaa  | tagctcttct  | gtgcaagatt  | acaatctaaa  | caattcggag  | 360  |
| aactcgacct  | tcctcctgag  | gcaaggacca  | cagccaactt  | cctcttacia  | gccgcacgca  | 420  |
| ttttgtcctt  | cagaaataga  | aataagaatg  | cttgctaaaa  | attatatattt | taccaataag  | 480  |
| accaatccaa  | taggttagatt | attagttact  | atgttaagaa  | atgaatcatt  | atcttttagt  | 540  |
| actattttta  | ctcaaattca  | gaagttagaa  | atgggaatag  | aaaatagaaa  | gagacgctca  | 600  |
| acctcaattg  | aagaacaggt  | gcaaggacta  | ttgaccacag  | gcctagaagt  | aaaaaaggga  | 660  |
| aaaaagagtg  | tttttgtcaa  | aataggagac  | aggtggtggc  | aaccagggac  | ttatagggga  | 720  |
| ccttacatct  | acagaccaac  | agatgcccc   | ttaccatata  | caggaagata  | tgacttaaat  | 780  |
| tgggataggt  | gggttacagt  | caatggctat  | aaagtgttat  | atagatccct  | cccttttctg  | 840  |
| gaaagactcg  | ccagagctag  | acctccttgg  | tgtatgttgt  | ctcaagaaga  | aaaagacgac  | 900  |
| atgaaacaac  | aggtacatga  | ttatatattt  | ctaggaacag  | gaatgcactt  | ttggggaaag  | 960  |
| attttccata  | ccaaggaggg  | gacagtggct  | ggactaatag  | aacattattc  | tgcaaaaact  | 1020 |
| catggcatga  | gttattatga  | atagccttta  | ttggcccaac  | cttgcggttc  | ccagggctta  | 1080 |
| agtaagtttt  | tgggtacaaa  | ctgttcttaa  | aacgaggatg  | tgagacaagt  | ggtttctctga | 1140 |
| cttggttttg  | tatcaaaggt  | tctgatctga  | gctctgagtg  | ttctattttc  | ctatgttctt  | 1200 |
| ttggaattta  | tcocaaatctt | atgtaaatgc  | ttatgtaaac  | caagatatata | aagagtgtctg | 1260 |
| attttttgag  | taaaacttgca | acagtcctaa  | cattcacctc  | ttgtgtgttt  | gtgtctgttc  | 1320 |
| gccatcccg   | ctccgctcgt  | cacttatcct  | tcactttcca  | gaggtcccc   | ccgcagaccc  | 1380 |
| cggcgacct   | caggctcgcc  | gactgcggca  | gctggcgccc  | gaacagggac  | cctcgataa   | 1440 |
| gtgaccttg   | tctctatttc  | tactatttgg  | tggttctctt  | gtattgtctc  | tttcttgtct  | 1500 |
| ggctatcatc  | acaagagcgg  | aacggactca  | ccatagggac  | caagctagcg  | cttctcgctg  | 1560 |
| cgtccaagac  | cctcaaagat  | ttttggcact  | tcgttgagcg  | aggcgatatc  | aggtatgaca  | 1620 |
| gcgcctcgcc  | gcaaggccag  | ctgcttgtcc  | gctcggtgc   | gggtggcacg  | gcaggatagg  | 1680 |
| ggatctctgc  | agttttggaa  | aaagatgtga  | taggtggcaa  | gcacctctgg  | cacggcaaat  | 1740 |
| acggggtaga  | agttgagggc  | cgggttgggc  | tcgcatgtgc  | cgttttcttg  | gcgtttgggg  | 1800 |
| ggtacgcgcg  | gtgagaatag  | gtggcgttcg  | taggcaaggc  | tgacatccgc  | tatggcgagg  | 1860 |
| ggcacatcgc  | tgcgctcttg  | caacgcgtcg  | cagataatgg  | cgcactggcg  | ctgcagatgc  | 1920 |
| ttcaacagca  | cgctgtctcc  | cacatctagg  | tagtcgccat  | gcctttcgtc  | ccccgcctcg  | 1980 |
| acttggtcct  | cgtttgcttc  | tgcggtgtcc  | tggtcttgct  | ttttatccct  | tggtgtgtact | 2040 |
| gagcggctcct | cgctcgtcttc | gcttacaaaa  | cctgggtcct  | gctcgataat  | cacttccctcc | 2100 |
| tcctcaagcg  | ggggtgcttc  | gacggggaag  | gtggtaggcg  | cggtggcggc  | atcggtggag  | 2160 |
| gcgggtgggtg | cgaactcaga  | gggggcgggt  | aggctgtcct  | tcttctcgac  | tgactccatg  | 2220 |
| atctttttct  | gcctatagga  | gaaggaaatg  | gccagtcggg  | aagaggagca  | gcgcgaaacc  | 2280 |
| acccccgagc  | gcggacgcgg  | tgccggcgca  | cgtcccccaa  | ccatggagga  | cgtgtcgtcc  | 2340 |
| ccgtccccgt  | cgccgcggcc  | tcgccgggcg  | cccccaaaaa  | agcggatgag  | gcggcgctac  | 2400 |
| gagtcogagg  | acgaggaaga  | ctcatcacia  | gacgcgctgg  | tgccgcgcac  | acccagcccg  | 2460 |
| cggccatcga  | cctcgccggc  | ggatttggcc  | attcgcccca  | agaagaaaaa  | gaagcgccct  | 2520 |
| tcctcccaagc | ccgagcgccc  | gccatcacca  | gaggtaatcg  | tggaacagca  | ggaagaaaga  | 2580 |
| gaagatgtgg  | cgctacaaat  | ggtgggtttc  | agcaaccac   | cggtgcta    | caagcatggc  | 2640 |
| aaaggaggta  | agcgcacagt  | gcggcggtcg  | aatgaagacg  | acccagtggc  | gcgtggatg   | 2700 |
| cggacgcaag  | aggaagagga  | agagcccagc  | gaagcggaaa  | gtgaaattac  | ggtgatgaac  | 2760 |
| ccgctgagtg  | tgccgatcgt  | gtctgcgtgg  | gagaagggca  | tggaaggctgc | gcgcgcgtg   | 2820 |
| atggacaagt  | accacgtgga  | taacgatcta  | aaggcgaaact | tcaaactact  | gcctgaccaa  | 2880 |
| gtggaagtct  | tggcgccgt   | atgcaagacc  | tggtgaacg   | aggagcaccg  | cgggttgacg  | 2940 |
| ctgaccttca  | tgagcaacaa  | gacctttgtg  | acgatgatgg  | ggcgattcct  | gcaggcgctac | 3000 |
| ctgcagtcgt  | ttgcagaggt  | gacctacaag  | catcacgagc  | ccacgggctg  | cgcgttgtgg  | 3060 |
| ctgcaccgct  | gcgctgagat  | cgaaggcgag  | cttaagtgtc  | tacacggaag  | cattatgata  | 3120 |
| aataaggagc  | acgtgattga  | aatggatgtg  | acgagcgaaa  | acgggcagcg  | cgcgctgaag  | 3180 |
| gagcagtcta  | gcaaggccaa  | gatcgtgaag  | aaccgggtggg | gccgaaatgt  | ggtgcagatc  | 3240 |
| tccaacaccg  | acgcaagggtg | ctgcgtgcac  | gacgcggcct  | gtccggccaa  | tcagttttcc  | 3300 |
| ggcaagtctt  | gcggcatgtt  | cttctctgaa  | ggcgcaaaagg | ctcaggtggc  | ttttaagcag  | 3360 |
| atcaaggctt  | ttatgcaggc  | gctgtatcct  | aacgcccaga  | ccgggcacgg  | tcaccttttg  | 3420 |
| atgccactac  | ggtgcgagtg  | caactcaaag  | cctgggcacg  | cgcccttttt  | gggaaggcag  | 3480 |
| ctaccaaagt  | tgactccgtt  | cgccctgagc  | aacgcggagg  | acctggacgc  | ggatctgatc  | 3540 |
| tccgacaaga  | gcgtgctggc  | cagcgtgcac  | caccggcgcc  | tgatagtgtt  | ccagtgtctg  | 3600 |
| aaccctgtgt  | atcgcaactc  | gcgcgcgcag  | ggcggaggcc  | ccaactgcga  | cttcaagata  | 3660 |
| tccggcgccc  | acctgctaaa  | cgcggttggtg | atgggtgcga  | gcctgtggag  | tgaaaacttc  | 3720 |
| accgagctgc  | cgcggatggg  | tgtgcctgag  | tttaagtggg  | gcactaaaca  | ccagtatcgc  | 3780 |
| aacgtgtccc  | tgccagtggc  | gcatagcgat  | gcgcggcaga  | acccctttga  | tttttaaacg  | 3840 |



|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gcgcagacgg  | caaggggtggg | ggtaaataat  | cacccgagag  | tgtacaaata  | aaagcatttg  | 3900 |
| cctttattga  | aagtgtctct  | agtacattat  | ttttacatgt  | ttttcaagt   | acaaaaagaa  | 3960 |
| gtggcgctcc  | taatctgcgc  | actgtggctg  | cggaagttagg | gcgagtggcg  | ctccaggaag  | 4020 |
| ctgtagagct  | gttcctgggt  | gcgacgcagg  | gtgggctgta  | cctggggact  | gttgagcatg  | 4080 |
| gagttgggta  | ccccggtaat  | aagggtcatg  | gtgggggtgt  | gatccatggg  | agtttggggc  | 4140 |
| cagttggcaa  | aggcgtggag  | aaacatgcag  | cagaatagtc  | cacaggcgcc  | cgagttgggc  | 4200 |
| ccctgtacgc  | tttgggtgga  | ctttccagc   | gttatacagc  | ggtcggggga  | agaagcaatg  | 4260 |
| gcgctacggc  | gcaggagtga  | ctcgtactca  | aactggtaaa  | cctgcttgag  | tcgctgggtca | 4320 |
| gaaaagccaa  | agggctcaaa  | gaggtagcat  | gtttttgagt  | gcgggttcca  | ggcaaaaggcc | 4380 |
| atccagtgta  | cgccccaggt  | ctcgcgacgc  | gccgtattga  | ctatggcgca  | ggcgagcttg  | 4440 |
| tgtggagaaa  | caaagcctgg  | aaagcgcttg  | tcatagggtg  | ccaaaaaata  | tggcccaaaa  | 4500 |
| ccaagatctt  | tgacaatggc  | tttcagttcc  | tgctcactgg  | agcccatggc  | ggcagctggt  | 4560 |
| gttgatgttg  | cttgcttctt  | ctgtgtgtgg  | cgttgccggc  | cgagaagggc  | gtgcgcagg   | 4620 |
| acacgggttc  | gatgacgcgc  | cggctcgccc  | gggtgcacag  | gaccacgtca  | aagacttcaa  | 4680 |
| acaaaacata  | aagaagggtg  | ggctcgtcca  | tgggatccat  | atataggggc  | cgggttataa  | 4740 |
| ttacctcagg  | tcgacctcga  | gggatctttg  | tgaaggaaac  | ttacttctgt  | ggtgtgacat  | 4800 |
| aattggacaa  | actacctaca  | gagatttaaa  | gctctaaggt  | aaatataaaa  | tttttaagtg  | 4860 |
| tataatgtgt  | taaactactg  | attctaattg  | ttgtgtatt   | ttagattcca  | acctatggaa  | 4920 |
| ctgatgaatg  | ggagcagtg   | tggaatgcct  | ttaatgagga  | aaacctgttt  | tgctcagaag  | 4980 |
| aatgccatc   | tagtgatgat  | gaggctactg  | ctgactctca  | acattctact  | cctccaaaaa  | 5040 |
| agaagagaaa  | ggtagaagac  | cccaaggact  | ttccttcaga  | attgctaagt  | tttttgagtc  | 5100 |
| atgctgtgtt  | tagtaataga  | actcttgctt  | gctttgctat  | ttacaccaca  | aaggaaaaag  | 5160 |
| ctgactgct   | atacaagaaa  | attatggaaa  | aatattctgt  | aacctttata  | agtaggcata  | 5220 |
| acagttataa  | tcataacata  | ctgttttttc  | ttactccaca  | caggcataga  | gtgtctgcta  | 5280 |
| ttaataacta  | tgctcaaaaa  | ttgtgtacct  | ttagcttttt  | aatgtgtaaa  | gggggttaata | 5340 |
| aggaatattt  | gatgtatagt  | gccttgacta  | gagatcataa  | tcagccatac  | cacatttgta  | 5400 |
| gaggttttac  | ttgctttaaa  | aaacctccca  | cacctccccc  | tgaacctgaa  | acataaaaatg | 5460 |
| aatgcaattg  | ttgttgttaa  | cttgtttatt  | gcagcttata  | atgggtacaa  | ataaagcaat  | 5520 |
| agcatcacaa  | atttcacaaa  | ttaaagcattt | ttttcactgc  | attctagtgt  | tggtttgtcc  | 5580 |
| aaactcatca  | atgtatctta  | tcatgtctgg  | atccggctgt  | ggaatgtgtg  | tcagttaggg  | 5640 |
| tgtggaaagt  | ccccaggctc  | cccagcaggc  | agaagtatgc  | aaagcatgca  | tctcaattag  | 5700 |
| tcagcaacca  | ggtgtggaaa  | gtcccaggc   | tcoccagcag  | gcagaagtat  | gcaaagcatg  | 5760 |
| catctcaatt  | agtcagcaac  | catagtcccg  | cccctaactc  | cgcccatccc  | gcccctaact  | 5820 |
| ccgcccagtt  | ccgcccattc  | tcgcgcccat  | ggctgactaa  | ttttttttat  | ttatgcagag  | 5880 |
| gccgaggccg  | cctcggccctc | tgagctattc  | cagaagtagt  | gaggaggctt  | ttttggaggc  | 5940 |
| ctaggctttt  | gcaaaaagct  | tcacgctgcc  | gcaagcactc  | agggcgcaag  | ggctgctaaa  | 6000 |
| ggaagcggaa  | cacgtagaaa  | gccagtcgct  | agaaacgggt  | ctgaccccg   | atgaatgtca  | 6060 |
| gctactgggc  | tatctggaca  | agggaaaaac  | caagcgcaaa  | gagaaagcag  | gtagcttgca  | 6120 |
| gtgggcttac  | atggcgatag  | ctagactggg  | cgggttttatg | gacagcaaac  | gaaccggaat  | 6180 |
| tgccagctgg  | ggcgccctct  | ggtaagggtg  | ggaagccctg  | caaagtaaac  | tggatggctt  | 6240 |
| tcttgccgcc  | aaggatctga  | tggcgaggg   | gatcaagatc  | tgatcaagag  | acaggatgag  | 6300 |
| gatcggttcg  | catgattgaa  | caagatggat  | tgacgcagcg  | ttctccggcc  | gcttgggtgg  | 6360 |
| agaggctatt  | cggctatgac  | tgggcacaa   | agacaatcgg  | ctgctctgat  | gccgcccgtg  | 6420 |
| tcgggtctgc  | agcgcagggg  | cgcccgggtc  | tttttgtcaa  | gaccgacctg  | tcgggtgccc  | 6480 |
| tgaatgaact  | gcaggacgag  | gcagcgcggc  | tatcggtggc  | ggccaacgac  | ggcggttccct | 6540 |
| gcgcagctgt  | gctcgacgtt  | gtcactgaag  | cgggaaggga  | ctggcttgcta | ttgggcgaag  | 6600 |
| tgccggggca  | ggatctcctg  | tcactctacc  | ttgctctgc   | cgagaaagta  | tcctcatgag  | 6660 |
| ctgatgcaat  | gcggcggtcg  | catacgcttg  | atccggctac  | ctgcccattc  | gaccaccaag  | 6720 |
| cgaaacatcg  | catcgagcga  | gcacgtactc  | ggatggaagc  | cggctctgtc  | gatcaggatg  | 6780 |
| atctggacga  | agagcatcag  | gggctcgcgc  | cagccgaact  | gttcgccagg  | ctcaaggcgc  | 6840 |
| gcagccccga  | cggcgaggat  | ctcgtcgtga  | cccatggcga  | tgctctgttg  | ccgaatatca  | 6900 |
| tgggtggaaaa | tggccgcttt  | tctggattca  | tcgactgtgg  | ccggctgggt  | gtggcggacc  | 6960 |
| gctatcagga  | catagcggtg  | gctaccgctg  | atatgtctga  | agagcttggc  | ggcgaatggg  | 7020 |
| ctgaccgctt  | cctcggtgctt | tacggtatcg  | ccgctcccga  | ttcgcagcgc  | atcgcttctt  | 7080 |
| atcgcttctt  | tgacgagttc  | ttctgagcgg  | gactctgggg  | ttcgaataga  | ccgaccaagc  | 7140 |
| gacgcccac   | ctgcccacac  | gagatttcga  | ttccaccgcc  | gccttctatg  | aaaggttggg  | 7200 |
| cttcgggaatc | gttttcgggg  | acgcgggctg  | gatgatcttc  | cagcgggggg  | atctcatgct  | 7260 |
| ggagttcttc  | gcccaacccg  | ggctcgatcc  | cctcgcgagt  | tggttcagct  | gctgcctgag  | 7320 |
| gctggacgac  | ctcgcggagt  | tctaccggca  | gtgcaaatcc  | gtcggcatcc  | aggaaaccag  | 7380 |
| cagcggtctat | ccgcgcaccc  | atgcccccca  | actgcaggag  | tggggaggca  | cgatggccgc  | 7440 |
| tttgggtccc  | gatctttgtg  | aaggaacctt  | acttctgtgg  | tgtgacataa  | ttggacaaa   | 7500 |
| tacctacaga  | gattttaaagc | tctaaggtaa  | atataaaatt  | tttaagtgtg  | taatgtgtta  | 7560 |



aactactgat tctaattggt tgtgtatntt agattccaac ctatggaact gatgaatggg 7620  
agcagtgggt gaatgccttt aatgagggaaa acctgttttg ctcagaagaa atgccatcta 7680  
gtgatgatga ggctactgct gactctcaac attctactcc tccaaaaaag aagagaaagg 7740  
tagaagaccc caaggacttt ccttcagaat tgctaagtnt tttgagtcac gctgtgttta 7800  
gtaatagaac tcttgcttgc ttttgctatnt acaccacaaa ggaaaaagct gcaactgctat 7860  
acaagaaaaat tatggaaaaa tttctgtaa cctttataag taggcataac agttataatc 7920  
ataacatact gtttttttctt actccacaca ggcatagagt gtctgctatt aataactatg 7980  
ctcaaaaatt gtgtaccttt agcttttttaa tttgtaaaagg ggttaataag gaatatntga 8040  
tgtatagtgc cttgactaga gatcataatc agccatacca ctttgtaga ggttttactt 8100  
gctttaaaaa acctcccaca cctcccctg aacctgaaac ataaaatgaa tgcaattggt 8160  
gttgtaact tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaaat 8220  
ttcacaaata aagcattttt ttcactgcac cctcaggaag cttatcctgtg gtttgtccaa actcatcaat 8280  
gtatcttatc atgtctggat ccccatagag ctccctctgtg tccctcataaa cctcaacctc 8340  
ctctacttga gaggacattc caatcatagg ctgcccaccc acctctgtg tccctctgtt 8400  
aattagggtca cttacacaaa aggaaattgg gtaggggttt ttcacagacc gctttctaa 8460  
ggtaatttta aaatatctgg gaagtccctt ccactgctgt gttccagaag gtttggtaaa 8520  
cagcccacaa atgtcaacag cagaacata caagctgtca gctttgcaca agggcccaac 8580  
acctgctca tcaagaagca ctgtggttgc tgtgttagta atgtgcaaaa caggaggcac 8640  
attttcccca cctgtgtagg ttccaaaata tctagtgttt tcatttttac ttggatcagg 8700  
aaccagcac tccactggat aagcattatc cttatccaaa acagcctgtg ggtcagtgtt 8760  
catctgctga ctgtcaactg tagcattttt tgggggttaca gtttgagcag gatatttggg 8820  
cctgtagtnt gctaacacac cctgcagctc caaagggtcc ccaccaacag caaaaaatg 8880  
aaaatttgac ccttgaatgg gttttccagc accatttttc tgagtntttt gttgcctga 8940  
atgcaaggnt aacatagcag ttaccccaat aacctcagtt ttaacagtaa cagcttccca 9000  
catcaaaata tttccacagg ttaagtctc atttaaatta ggcaaaaggaa ttcttgaaga 9060  
cgaaagggtc tctgtatagc cctattttta taggttaatg tcatgataat aatgggttct 9120  
tagacgtcag gtggcacttt tccgctcatg agacaataac cccctatttg tttatttttc 9180  
taaatacatt caaatatgta gagtattcaa catttccgtg cctgataaat gcttcaataa 9240  
tattgaaaaa ggaagagtat gtttgcac ccagaaaacgc tgggtgaaagt tccctttttt 9300  
gcggcatttt gccttccgtg agtgggttac agaaactgg atctcaacag aaaagatgct 9360  
gaagatcagt tgggtgcacg agaacgtttt ccaatgatga gcaactttta cggtaagatc 9420  
cttgagagtt ttgccccga tgttgacgoc gggcaagagc aactcggtcg agttctgcta 9480  
tgtggcgcggt tattatcccc atgacttggg cagtcacag gctgataacac cgcatacac 9540  
tattctcaga atgacttggg cagtgtctgc agtaaacac ctaacatgga caaatgggg 9600  
atgacagtta caacgatcgg aggaccgaag cagtaacac gtgataacac tgccgccaac 9660  
ttacttctga ctcgccttga tctgtgggaa ccggagctga atgaagccat accaaacgac 9720  
gatcatgtaa cccagatgcc tgcagcaact ttaatagact ggatggaggc ggataaagt 9780  
gagcgtgaca ccactagctt cggcaacaa gctggctggg ttattgctga taaatctgga 9840  
gaactactta ttctgcgtc cggccttccg cggatcatt taagccctcc 10020  
gcaggaccac gtgggtctcg gacggtgag agcagactgg ggcagatgg aaatagacag 10080  
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 10140  
atcgtctgaga taggtgcctc actgattaag catttgtaac tgctagacca agtttactca 10200  
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 10260  
ctttttgata atctcatgac caaaatccct tttocttcca ctgagcgtca cgtaatctgc 10320  
gaccccgtag aaaagatcaa aggatcttct tttttctgcy gtttgccgga tcaagagcta 10380  
tgcttgcaaa caaaaaaacc accgctacca gcggtggtt agataccaaa tactgtcctt 10440  
ccaactcttt ttccgaaggt aactggctt agcagagcgc tagcaccgoc tacatacctc 10500  
ctagtgtagc cgtagttagg ccaccacttc agtggtgct gccagtggcg ataagtctg 10560  
gctctgctaa tctgtttacc accggataag gcgagcggt cggggtgaac ggggggttcg 10620  
ttggactcaa gacgatagtt gcgaacgacc tacacgaac tgagatacct acagcgtgag 10680  
tgcacacagc ccagcttggg gcccgaagg cagagggag cttccagggg acaggtatcc ggttaagcggc 10740  
ctatgagaaa gcgccacgct cagagggag cctctgact gagcgtcgat ttttgtgat gttatctttat 10800  
agggtcgga ggtttcgcca cctctgact gcgcagcaac taccgctctg atctgtgag ctcgtcagg 10860  
agtcctgtcg tatggaaaaa ctttctgcg taccgctcgc cggatnttct taccgtatt 10920  
gggaggagcc ctcacatggt agtgagctga ggcgtctgag cactctcagt ctgtgctgag 10980  
tggtcttttg accgctttg aagcggaaga cctatggtg tccctgcttg cagcgagtca 11040  
accgcttttg atttacacc cagtatctgc tccctgcttg gctacaacaa ggcaaggctt ggttaggcgt 11100  
gtgagcgagg atttacacc cagtatctgc tccctgcttg gctacaacaa ggcaaggctt ggttaggcgt 11160  
atttacacc cagtatctgc tccctgcttg gctacaacaa ggcaaggctt ggttaggcgt 11220  
gctacaacaa ggcaaggctt gctacaacaa ggcaaggctt gctacaacaa ggcaaggctt 11280

```

tttgcgctgc ttcgcgatgt acgggccaga tatacgcgta tctgagggga ctaggggtgtg 11340
tttaggcgaa aagcggggct tcggttgtag gcggttagga gtcccctcag gatatagtag 11400
tttcgctttt gcatagggag ggggaaatgt agtcttatgc aatacacttg tagtcttgca 11460
acatggtaac gatgagttag caacatgcct tacaaggaga gaaaaagcac cgtgcatgcc 11520
gattgggtgga agtaagggtg tacgatcgtg ccttattagg aaggcaacag acgggtctga 11580
catggattgg acgaaccact

```

```

<210> 36
<211> 53
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: primer

```

```

<400> 36
gtcactcgag gactcggctg actgaaaatg agacatatta tctgccacgg acc 53

```

```

<210> 37
<211> 36
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: primer

```

```

<400> 37
cgagatcgat cacctccggt acaagggttg gcatag 36

```

```

<210> 38
<211> 37
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: primer

```

```

<400> 38
catgaagatc tggaagggtg tgagggtacga tgagacc 37

```

```

<210> 39
<211> 51
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: primer

```

```

<400> 39
gcgacttaag cagtcagctg agacagcaag acacttgctt gatccaaatc c 51

```

```

<210> 40
<211> 38
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: primer

```

```

<400> 40
cacgaattcg tcagcgcttc tcgtcgcgtc caagaccc 38

```

<210> 41  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 41  
caccgccggg aggcggcggc gacggggacg gg

32

<210> 42  
<211> 7231  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: plasmid pDV80

<400> 42  
ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg cgagcaaaat ttaagctaca 60  
acaaggcaag gcttgaccga caattgcatg aagaatctgc ttagggtagt gcgttttgcg 120  
ctgcttcgcg atgtacgggc cagatatagc cgttgacatt gattattgac tagttattaa 180  
tagtaataca ttacggggtc attagttcat agcccatata tggagttccg cggtacataa 240  
cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 300  
atgacgtatg ttcccatagt aacgcccaata gggactttcc attgacgtca atgggtggac 360  
tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 420  
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 480  
tgggactttc ctacttggca gtacatctac gtattagtc tgcgtattac catgggtgatg 540  
cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 600  
ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 660  
aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt acggtgggag 720  
gtctatataa gcagagctct ctggctaact agagaaccca ctgcttactg gcttatcgaa 780  
attaatacga ctccactatg ggagacccaa gctggctagc gtttaaactt aagcttggtg 840  
ccgagctcgg atccactctc ttccgcatcg ctgtctgcga gggccagctg ttggggtgag 900  
tactccctct gaaaagcggg catgacttct gcgctaagat tgctcagttt caaaaacgag 960  
gaggatttga tattcacctg gcccgcggtg atgcctttga ggggtggccgc atccatctgg 1020  
tcagaaaaga caatcttttt gttgtcaagc gctgtgcaa acgaccgta gagggcgttg 1080  
gacagcaact tggcgatgga gcgcagggtt tgggttttgt cgcgatcggc gcgctccttg 1140  
gccgcgatgt ttagctgcac gtattcgcgc gcaacgcacc gccattcggg aaagacgggtg 1200  
gtgcgctcgt cgggcaccag gtgcacgcgc caaccgcggt tgtgcagggt gacaaggcca 1260  
acgctgggtg ctacctctcc gcgtaggcgc tcgttgggcc agcagaggcg gccgcccttg 1320  
cgcgagcaga atggcggtag ggggtctagc tgcgtctcgt ccgggggggtc tgcgtccacg 1380  
gtaaagaccc cgggcagcag gcgcgcgctg aagtagtcta tcttgcatec ttgcaagtct 1440  
agcgctgct gccatgcgcg ggcggcaagc gcgcgctcgt atgggttgag tgggggaccc 1500  
catggcatgg ggtgggtgag cgcggaggcg tacatgccgc aaatgtcgta aacgtagagg 1560  
ggctctctga gtattccaag atatgtaggg tagcatcttc caccgcggat gctggcgcg 1620  
acgtaatcgt atagttcgtg cgaggagcg aggaggtcgg gaccgaggtt gctacggggc 1680  
ggctgctctg ctcggaagac tatctgcctg aagatggcat gtgagttgga tgatatgggt 1740  
ggacgctgga agacgttgaa gctggcgtct gtgagacctc ccgcgtcacg cacgaaggag 1800  
gcgtaggagt cgcgcagctt gttgaccagc tcggcggtga cctgcacgtc tagggcgcg 1860  
tagtccaggg tttccttgat gatgtcatac ttatcctgtc cctttttttt ccacagctcg 1920  
cggttgagga caaactcttc gcggtctttc cagtactctt ggatcggaac cccgtcggcc 1980  
tccgaacgag atccgtactc cgccgcccag ggacctgagc gagtccgcac cgaccggatc 2040  
ggaaaacctc tcgagaaagg cgtctaacca gtcacagtcg caagatccaa gatgaagcgc 2100  
gcccccccca gcgaagatga cttcaacccc ctctaccctc atggctacgc gcggaatcag 2160  
aatatccctt tccctactcc cccctttgtc tctccgatg gattcaaaaaa cttccccct 2220  
ggggtactgt cactcaaaact ggctgatcca atcaccatta ccaatgggga tgtatccctc 2280  
aaggtgggag gtggtctcac tttgcaagat ggaagcctaa ctgtaaacc taaggctcca 2340  
ctgcaagtta atactgataa aaaacttgag cttgcatatg ataaccatt tgaaagtagt 2400  
gctaataaac ttagttttaa agtaggacat ggattaaaag tattagatga aaaaagtgtc 2460

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gcggggttaa  | aagatttaat  | tggcaaaactt | gtgggttttaa | caggaaaagg  | aataggcact  | 2520 |
| gaaaatttag  | aaaatacaga  | tggttagcagc | agaggaattg  | gtataaatgt  | aagagcaaga  | 2580 |
| gaagggttga  | catttgacaa  | tgatggatac  | ttggtagcat  | ggaacccaaa  | gtatgacacg  | 2640 |
| cgcacacttt  | ggacaacacc  | agacacatct  | ccaaactgca  | caattgctca  | agataaggac  | 2700 |
| tctaaactca  | ctttgggtact | tacaaagtgt  | ggaagtcaaa  | tattagctaa  | tgtgtctttg  | 2760 |
| atttgtgtcg  | caggaaagta  | ccacatcata  | aataataaga  | caaataccaa  | aataaaaaagt | 2820 |
| tttactatta  | aactgctatt  | taataagaac  | ggagtgcctt  | tagacaactc  | aaatcttgga  | 2880 |
| aaagcttatt  | ggaacttttag | aagtggaaat  | tccaatgttt  | cgacagctta  | tgaaaaagca  | 2940 |
| atttggtttta | tgccctaattt | ggtagcgtat  | ccaaaaccca  | gtaattctaa  | aaaatatgca  | 3000 |
| agagacatag  | tttatgggaac | tatatatcct  | ggtaggaaaac | ctgatcagcc  | agcagtcatt  | 3060 |
| aaaactacct  | ttaaccaaga  | aactggatgt  | gaatactcta  | tcacatttaa  | ctttagttgg  | 3120 |
| tccaaaacct  | atgaaaatgt  | tgaatttgaa  | accacctctt  | ttaccttctc  | ctatatgtcc  | 3180 |
| caagaatgaa  | agagcgggccg | ctcgagtcta  | gagggcccggt | ttaaacccgc  | tgatcagcct  | 3240 |
| cgactgtgcc  | ttctagttagc | cagccatctg  | ttgtttggcc  | ctcccccggtg | ccttccttga  | 3300 |
| ccctggaagg  | tgccactccc  | actgtccttt  | cctaataaaa  | tgaggaaatt  | gcatcgcatt  | 3360 |
| gtctgagtag  | gtgtcattct  | attctggggg  | gtgggggtggg | gcaggacagc  | aagggggagg  | 3420 |
| attgggaaga  | caatagcagg  | catgctgggg  | atgctgtggg  | ctctatggct  | tctgaggcgg  | 3480 |
| aaagaaccag  | ctgggggtct  | aggggggtatc | cccacgcgcc  | ctgtagcgcc  | gcattaagcg  | 3540 |
| cgggcggtgt  | gggtggttacg | cgacgctga   | ccgtacact   | tgccagcgcc  | ctagcgcccg  | 3600 |
| ctcctttcgc  | ttctttccct  | tcctttctcg  | ccacgttcgc  | cggtttccc   | cgtaagctc   | 3660 |
| taaatcgggg  | catcccttta  | gggttcogat  | ttagtgtctt  | acggcacctc  | gaccccaaaa  | 3720 |
| aaactgatta  | gggtgatggg  | tcacgtagt   | ggccatcgcc  | ctgatagacg  | gtttttcgcc  | 3780 |
| ctttgacggt  | ggagtccacg  | ttctttaata  | gtggactctt  | gttccaaact  | ggaacaacac  | 3840 |
| tcaaccctat  | ctcggtctat  | tcctttgatt  | tataagggat  | tttggggatt  | tcggcctatt  | 3900 |
| ggttaaaaaa  | tgagctgatt  | taacaaaaat  | ttaacgcgaa  | ttaattctgt  | ggaatgtgtg  | 3960 |
| tcagttaggg  | tgtggaaagt  | ccccaggctc  | cccaggcagg  | cagaagtatg  | caaagcatgc  | 4020 |
| atctcaatta  | gtcagcaacc  | aggtgtggaa  | agtcgccagg  | ctccccagca  | ggcagaagta  | 4080 |
| tgcaaagcat  | gcatctcaat  | tagtcagcaa  | ccatagtccc  | gcccctaact  | ccgcccattc  | 4140 |
| cgccctaac   | tcgccccagt  | tcgccccatt  | ctccgcccga  | tggtgacta   | atttttttta  | 4200 |
| tttatgcaga  | ggccgaggcc  | gcctctgcct  | ctgagctatt  | ccagaagtag  | tgaggagggt  | 4260 |
| tttttggagg  | cctaggcttt  | tgcaaaaagc  | tcccgggagc  | ttgtatatcc  | attttcggat  | 4320 |
| ctgatcagca  | cgtgttgaca  | attaatcatc  | ggcatagtat  | atcggcatag  | tataatacga  | 4380 |
| caaggtgagg  | aactaaacca  | tggccaagtt  | gaccagtgc   | gttccgggtg  | tcaccgcgcg  | 4440 |
| cgacgtcgcc  | ggagcgggtc  | agttctggac  | cgaccggctc  | gggttctccc  | gggacttcgt  | 4500 |
| ggaggacgac  | ttcgccgggtg | tggtccggga  | cgacgtgacc  | ctgttcatca  | gcgcgggtcca | 4560 |
| ggaccagggtg | gtgcccggaca | acaccctggc  | ctgggtgtgg  | gtgcccggcc  | tgagcagagt  | 4620 |
| gtacgccgag  | tggtcggagg  | tcgtgtccac  | gaacttccgg  | gacgcctccg  | ggccggccat  | 4680 |
| gaccgagatc  | ggcgagcagc  | cgtgggggcg  | ggagtccgc   | ctgcgcgacc  | cggccggcaa  | 4740 |
| ctgcgtgcac  | ttcgtggcgg  | aggagcagga  | ctgacacgtg  | ctacgagatt  | tcgattccac  | 4800 |
| cgccgccttc  | tatgaaaggt  | tgggcttcgg  | aatcgttttc  | cgggacgccc  | gctggatgat  | 4860 |
| cctccagcgc  | gggatctca   | tgctggagtt  | cttcgcccac  | cccaacttgt  | ttattgcagc  | 4920 |
| ttataatggt  | tacaaaataa  | gcaatagcat  | cacaaatttc  | acaaataaag  | catttttttc  | 4980 |
| actgcattct  | agttgtggtt  | tgtccaaact  | catcaatgta  | tcttatcatg  | tctgtatacc  | 5040 |
| gtcgacctct  | agctagagct  | tggcgtaatc  | atggctcatg  | ctgtttcctg  | tgtgaaattg  | 5100 |
| ttatccgctc  | acaattccac  | acaacatacg  | agccggaagc  | ataaagtgtg  | aagcctgggg  | 5160 |
| tgctaatga   | gtgagctaac  | tcacattaat  | tgcgttgccg  | tcactgccc   | ctttccagtc  | 5220 |
| gggaaacctg  | tcgtgccagc  | tgcattaatg  | aatcgcccaa  | cgcgcgggga  | gaggcggttt  | 5280 |
| gcgtattggg  | cgctcttcgg  | cttcctcgct  | caactgactc  | ctgcgctcgg  | tcgttcggct  | 5340 |
| gcggcgagcg  | gtatcagctc  | actcaaaggc  | ggtaatacgg  | ttatccacag  | aatcagggga  | 5400 |
| taacgcagga  | aagaacatgt  | gagcaaaaag  | ccagcaaaaag | gccaggaacc  | gtaaaaaggc  | 5460 |
| cgctgtgctg  | gcgtttttcc  | ataggctccg  | ccccctgac   | gagcatcaca  | aaaatcgacg  | 5520 |
| ctcaagtccg  | aggtggcgaa  | acccgacagg  | actataaaga  | taccaggcgt  | ttccccctgg  | 5580 |
| aagctccctc  | gtgcgctctc  | ctgttccgac  | cctgccgctt  | accggatacc  | tgtccgcctt  | 5640 |
| tctcccttcg  | ggaagcgtgg  | cgctttctca  | atgctcacgc  | tgtaggtatc  | tcagttcggt  | 5700 |
| gtaggctcgtt | cgctccaagc  | tgggctgtgt  | gcacgaaccc  | cccgttcagc  | ccgaccgctg  | 5760 |
| cgcttatccc  | ggtaactatc  | gtcttgagtc  | caacccggtg  | agacacgact  | tatcgccact  | 5820 |
| ggcagcagcc  | actggtaaca  | ggattagcag  | agcgagggtat | gtaggcggtg  | ctacagagtt  | 5880 |
| cttgaagtgg  | tggcctaact  | acggctacac  | tagaaggaca  | gtatttggta  | tctgcgctct  | 5940 |
| gctgaagcca  | ggttaccttcg | gaaaaagagt  | tggtgactct  | tgatccggca  | aacaaaccac  | 6000 |
| cgctggtagc  | ggtaggtttt  | ttgtttgcaa  | gcagcagatt  | acgcgcagaa  | aaaaaggatc  | 6060 |
| tcaagaagat  | cctttgatct  | tttctacggg  | gtctgacgct  | cagtggaaacg | aaaactcacg  | 6120 |
| ttaagggtatt | ttggctcatga | gattatcaaa  | aaggatcttc  | acctagatcc  | ttttaaatta  | 6180 |

```

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 6240
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 6300
ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtcg 6360
tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc 6420
agccggaagg gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca tccagtctat 6480
taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt 6540
tgccattgct acaggcatcg tgggtgtcacg ctctcggttt ggtatggctt cattcagctc 6600
cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcgggttag 6660
ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt 6720
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac 6780
tgggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgtctctg 6840
cccggcgctca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat 6900
tggaataacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc 6960
gatgtaaccc actcgtgcac ccaactgacg ttcagcatct tttactttca ccagcggttc 7020
tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa 7080
atgttgataa ctcatactct tcctttttca atattattga agcatttatc agggttattg 7140
tctcatgagc ggatacatat ttgaatgtat ttgaaaaaat aaacaaatag gggttccgcg 7200
cacatttccc cgaaaagtcg cacctgacgt c 7231

```

<210> 43  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 43  
 tgtcttggat ccaagatgaa gcgcgcccgc cccagcgaag atgacttc

48

<210> 44  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: primer

<400> 44  
 aaacacggcg gccgctcttt cattcttg

28

<210> 45  
 <211> 17  
 <212> PRT  
 <213> Ad 37 N-terminus

<400> 45  
 Met Ser Lys Arg Leu Arb Val Glu Asp Asp Phe Asn Pro Val Tyr Pro Tyr  
 1 5 10 15

<210> 46  
 <211> 6  
 <212> PRT  
 <213> artificial sequence

<400> 46  
 Lys Arg Ala Arg Pro Ser  
 1 5

<210> 47  
 <211> 17

<212> PRT  
<213> Ad5 modified N-terminus

<400> 47  
Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro Tyr  
1 5 10 15

<210> 48  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 48  
ggatccatgg gatacttggt agca 24

<210> 49  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 49  
gcaactcgag tcattcttgg gcaatatagg 30

<210> 50  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 50  
cgcgctgact cttaaggact agtttc 26

<210> 51  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 51  
gcgcttaatt aacatcatca ataataacc ttatattt 37